

# LOCAL CORRELATION ENTROPY

VLADIMÍR ŠPITALSKÝ

**ABSTRACT.** Local correlation entropy, introduced by Takens in 1983, represents the exponential decay rate of the relative frequency of recurrences in the trajectory of a point, as the embedding dimension grows to infinity. In this paper we study relationship between the supremum of local correlation entropies and the topological entropy. For dynamical systems on graphs we prove that the two quantities coincide. Moreover, there is an uncountable set of points with local correlation entropy arbitrarily close to the topological entropy. On the other hand, we construct a strictly ergodic subshift with positive topological entropy having all local correlation entropies equal to zero. As a necessary tool, we derive an expected relationship between the local correlation entropies of a system and those of its iterates.

## 1. INTRODUCTION

A (topological) dynamical system is a pair  $(X, f)$  where  $X$  is a compact metric space and  $f : X \rightarrow X$  is a continuous map. A point  $x \in X$  is recurrent when its trajectory  $(f^n(x))_{n=0}^\infty$  returns repeatedly to every neighborhood of  $x$ . The topological version of the famous Poincaré recurrence theorem states that, with respect to every invariant Borel measure, almost every point is recurrent. So if we look at the trajectory of a typical point  $x$ , we see infinitely many indices  $n$  such that  $f^n(x)$  is close to  $x$ . Moreover, continuity of  $f$  implies that we see infinitely many pairs of indices  $i \neq j$  such that  $f^i(x)$  is close to  $f^j(x)$ . Such pairs are called recurrences.

Recurrences can be effectively visualized via recurrence plots, introduced by Eckmann, Kamphorst, and Ruelle in [6]. In its basic form, a recurrence plot is a black-and-white square image with black pixels representing recurrences. Quantitative study of patterns occurring in recurrence plots is the subject of the so-called recurrence quantification analysis initiated by Zbilut and Weber [28]; for a recent survey see [16].

In connection with correlation dimension [9, 10] and correlation entropy [24] introduced in the beginning of 80's, the so-called correlation sums were studied. Recall that the *correlation sum*  $C_\varrho(x, n, \varepsilon)$  of (the beginning of) the trajectory of a point  $x$  is

$$(1.1) \quad C_\varrho(x, n, \varepsilon) = \frac{1}{n^2} \text{card} \{ (i, j) : 0 \leq i, j < n, \varrho(f^i(x), f^j(x)) \leq \varepsilon \},$$

where  $\varrho$  is the metric of  $X$ ,  $n \in \mathbb{N}$ , and  $\varepsilon > 0$ . It is simply the relative frequency of recurrences seen in the initial segment of the trajectory of  $x$ , with closeness defined by the metric  $\varrho$  and the distance threshold  $\varepsilon$  (with pairs  $(i, i)$  counted as recurrences). Correlation sums appear naturally in different contexts. They are used in the estimation of correlation dimension and correlation entropy. In the recurrent quantification analysis, several of the basic quantitative characteristics can be expressed in terms of correlation sums [11]. Also note that, by removing the diagonal pairs  $(i, i)$ , correlation sum becomes a  $U$ -statistic [5, 1].

One of the fundamental results states that, with respect to any  $f$ -ergodic measure  $\mu$ , correlation sums of  $\mu$ -almost every point  $x$  converges to the *correlation integral*

$$(1.2) \quad c_\varrho(\mu, \varepsilon) = \mu \times \mu \{ (y, z) \in X \times X : \varrho(y, z) \leq \varepsilon \} = \int_X \mu B_\varrho(x, \varepsilon) d\mu(x)$$

where  $B_\varrho(x, \varepsilon)$  denotes the closed ball with the center  $x$  and radius  $\varepsilon$ . This was proved (by different methods and under different conditions) in [19, 20, 1, 22, 15]. This result justifies the use of correlation sums in estimating the correlation dimension, as suggested by [9, 10].

---

2010 *Mathematics Subject Classification.* Primary 37B40, 28D20; Secondary 54H20.

*Key words and phrases.* local correlation entropy, topological entropy, strict ergodicity.

The correlation entropy, introduced by Takens in [24], is a quantitative measure based on correlation sums / integrals. To define it, in (1.1) and (1.2) replace the metric  $\varrho$  by Bowen's one

$$(1.3) \quad \varrho_m^f(y, z) = \max_{0 \leq i < m} \varrho(f^i(y), f^i(z)) \quad (y, z \in X).$$

The obtained quantities are the correlation sum  $C_m^f(x, n, \varepsilon)$  and the correlation integral  $c_m^f(\mu, \varepsilon)$  corresponding to the trajectory of  $x$  embedded to  $X^m$ . The *upper and lower correlation entropy of an  $f$ -invariant measure  $\mu$*  [2, p. 361] measures exponential decay rate of correlation integrals as  $m$  grows to infinity

$$(1.4) \quad \bar{h}_{\text{cor}}(f, \mu) = \lim_{\varepsilon \rightarrow 0} \limsup_{m \rightarrow \infty} (-1/m) \log c_m^f(\mu, \varepsilon), \quad h_{\text{cor}}(f, \mu) = \lim_{\varepsilon \rightarrow 0} \liminf_{m \rightarrow \infty} (-1/m) \log c_m^f(\mu, \varepsilon),$$

This entropy is a member of a 1-parameter family of entropies, see [25].

The definition above which is recently used in the literature, differs from the original one [24] by using correlation integrals instead of correlation sums. Consequently, it depends on an invariant measure  $\mu$  instead of a point  $x$ . To distinguish the original definition from the recently used one, the correlation entropy of  $f$  at a point  $x$  will be called local. So, following [24], the *upper and lower local correlation entropies of  $f$  at  $x$*  are defined by

$$(1.5) \quad \bar{h}_{\text{cor}}(f, x) = \lim_{\varepsilon \rightarrow 0} \limsup_{m \rightarrow \infty} (-1/m) \log c_m^f(x, \varepsilon), \quad h_{\text{cor}}(f, x) = \lim_{\varepsilon \rightarrow 0} \liminf_{m \rightarrow \infty} (-1/m) \log c_m^f(x, \varepsilon),$$

where

$$(1.6) \quad \bar{c}_m^f(x, \varepsilon) = \limsup_{n \rightarrow \infty} C_m^f(x, n, \varepsilon), \quad c_m^f(x, \varepsilon) = \liminf_{n \rightarrow \infty} C_m^f(x, n, \varepsilon).$$

(Note that, in [24], the author considered the lower entropy only.) Of course, due to the convergence of correlation sums to the correlation integral, these local correlation entropies are often equal to the correlation entropy of a measure  $\mu$ . Nevertheless, we believe that it deserves to study these local correlation entropies, for which we have several reasons. First, the ergodic results hold (usually) only for almost every point, but, from the topological point of view, local correlation entropy at every point should be considered. Second, since local correlation entropy depends solely on the trajectory of a selected point, it is computationally more tractable than correlation integral. In fact, when estimating correlation entropy of an invariant measure  $\mu$ , correlation sums are often used and thus the local correlation entropy is being estimated; see e.g. [2, §7.7]. Finally, study of local correlation entropies can yield new results, which have not yet been obtained for correlation entropy of a measure.

Let us now briefly outline the main results of this paper. We start with summarizing basic properties of the local correlation entropy. One of them is the relationship between local correlation entropies of  $f$  and those of its iterates  $f^k$ . Since we were not able to find a corresponding result in the literature, we included a proof of it in this paper. The proof is based on a combinatorial lemma (see §3.2), which gives a relationship between correlation sum of  $f$  at a point  $x$  and correlation sums of  $f^k$  at points  $f^h(x)$  ( $0 \leq h < k$ ), see Lemma 18.

**Theorem A.** *Let  $(X, f)$  be a dynamical system. Then, for every  $k \in \mathbb{N}$  and  $x \in X$ ,*

$$\bar{h}_{\text{cor}}(f^k, x) = k \cdot \bar{h}_{\text{cor}}(f, x), \quad h_{\text{cor}}(f^k, x) = k \cdot h_{\text{cor}}(f, x).$$

The basic motivation of the paper comes from studying the relationship between the local correlation entropies and the topological entropy of the system  $(X, f)$ . Already Takens [24] proved that the lower local correlation entropy is bounded from above by the topological entropy of  $f$  restricted to the orbit closure of  $x$ . In Proposition 21 we prove that this is true also for the upper local correlation entropy, which yields that

$$\sup_{x \in X} h_{\text{cor}}(f, x) \leq \sup_{x \in X} \bar{h}_{\text{cor}}(f, x) \leq h_{\text{top}}(f).$$

We will show that, for dynamical systems on graphs, the above inequalities are in fact equalities. Recall that a graph is a continuum which can be written as the union of finitely many arcs any two of which are either disjoint or intersect only in one or both of their end points.

**Theorem B.** *Let  $X$  be a graph and  $f : X \rightarrow X$  be a continuous map. Then*

$$\sup_{x \in X} h_{\text{cor}}(f, x) = \sup_{x \in X} \bar{h}_{\text{cor}}(f, x) = h_{\text{top}}(f).$$

Moreover, for every  $h < h_{\text{top}}(f)$  there is a Cantor set  $X_h \subseteq X$  such that  $h_{\text{cor}}(f, x) \geq h$  for every  $x \in X_h$ .

For general dynamical systems, the supremum of local correlation entropies can be strictly smaller than the topological entropy. We prove this by constructing a strictly ergodic subshift with positive entropy and with all local correlation entropies equal to zero; our construction is a modification of Grillenberger's one [12].

**Theorem C.** *There is a subshift  $(X, \sigma)$  such that*

- (a)  $(X, \sigma)$  is strictly ergodic;
- (b)  $(X, \sigma)$  has positive topological entropy;
- (c) the local correlation entropy  $h_{\text{cor}}(\sigma, y)$  at every  $y \in X$  is zero;
- (d) the correlation entropy  $h_{\text{cor}}(\sigma, \mu)$  of the unique invariant measure  $\mu$  is zero.

For some other results which are worth mentioning and are not covered by Theorems A–C, see Corollary 13 and Propositions 3 and 23.

The paper is organized as follows. In §2 we recall definitions and known facts which will be required later. In §§3 and 4 we prove Theorems A and B. A technical lemma concerning strictly ergodic subshifts is given in §5. Finally, in §6 we prove Theorem C.

## 2. PRELIMINARIES

We write  $\mathbb{N}$  ( $\mathbb{N}_0$ ) for the set of positive (nonnegative) integers. If no confusion can arise, segments of integers  $\{n, n+1, \dots, m-1\}$  ( $n < m$ ) will be denoted by  $[n, m)$ . For  $x \in \mathbb{R}$ ,  $\lceil x \rceil$  and  $\lfloor x \rfloor$  denotes the ceiling and the floor of  $x$ , that is, the smallest integer greater than or equal to  $x$ , and the largest integer smaller than or equal to  $x$ . The cardinality of a set  $A$  is denoted by  $|A|$  or by  $\text{card } A$ . By  $\log$  we mean the natural logarithm.

Let  $X = (X, \varrho)$  be a metric space and  $A$  be a subset of it. The diameter of a subset  $A$  of  $X$  is denoted by  $\text{diam}_\varrho(A)$ . By  $B_\varrho(x, \varepsilon)$  we mean the closed ball with the center  $x$  and radius  $\varepsilon$ , and by  $B_\varrho(A, \varepsilon)$  we mean the union of all  $B_\varrho(x, \varepsilon)$  with  $x \in A$ . The set  $A$  is called  $\varepsilon$ -separated if  $\varrho(x, y) > \varepsilon$  for every  $x \neq y$  from  $A$ . It is said to  $\varepsilon$ -span  $X$  if  $B_\varrho(A, \varepsilon) = X$ . The smallest cardinality of an  $\varepsilon$ -spanning subset of  $X$  is denoted by  $r_\varrho(\varepsilon, X)$ , and the largest cardinality of an  $\varepsilon$ -separated subset of  $X$  is denoted by  $s_\varrho(\varepsilon, X)$ . If  $X$  is compact, both  $r_\varrho(\varepsilon, X)$  and  $s_\varrho(\varepsilon, X)$  are always finite, and we can define the upper and lower box dimension of  $X$  [8, §3.1]

$$\bar{d}_{\text{box}}(X; \varrho) = \limsup_{\varepsilon \rightarrow 0} \frac{\log r_\varrho(\varepsilon, X)}{-\log \varepsilon} \quad \text{and} \quad \underline{d}_{\text{box}}(X; \varrho) = \liminf_{\varepsilon \rightarrow 0} \frac{\log r_\varrho(\varepsilon, X)}{-\log \varepsilon}.$$

A *measure-theoretical dynamical system* is a quadruple  $(X, \mathcal{F}, \mu, f)$ , where  $X$  is a nonempty set,  $\mathcal{F}$  is a  $\sigma$ -algebra of subsets of  $X$ ,  $\mu$  is a probability measure on  $(X, \mathcal{F})$ , and  $f : X \rightarrow X$  is an  $\mathcal{F}$ -measurable map preserving  $\mu$  (that is,  $\mu(f^{-1}(A)) = \mu(A)$  for every  $A \in \mathcal{F}$ ). The system  $(X, \mathcal{F}, \mu, f)$  is called *ergodic* if  $\mu(A) \in \{0, 1\}$  for every  $A \in \mathcal{F}$  such that  $f^{-1}(A) = A$ .

A *(topological) dynamical system* is a pair  $(X, f)$  where  $X = (X, \varrho)$  is a compact metric space and  $f : X \rightarrow X$  is a continuous map. A set  $A \subseteq X$  is said to be *f-invariant* if  $f(A) \subseteq A$ . A system  $(X, f)$  is *minimal* if there is no nonempty proper closed  $f$ -invariant subset of  $X$ . Every point of a minimal system  $(X, f)$  is *almost periodic*: for every neighborhood  $U$  of  $x$  the return time set  $N(x, U)$  is syndetic (that is, it has bounded gaps).

An *f-invariant measure* of  $(X, f)$  is any Borel probability measure  $\mu$  such that  $(X, \mathcal{B}, \mu, f)$ , with  $\mathcal{B}$  denoting the Borel  $\sigma$ -algebra on  $X$ , is a measure-theoretical dynamical system. If  $(X, \mathcal{B}, \mu, f)$  is ergodic we say that  $\mu$  is *f-ergodic*. A system  $(X, f)$  is called *uniquely ergodic* if it has unique invariant measure; if it is also minimal it is called *strictly ergodic*.

Let  $(X, f)$  be a (topological) dynamical system and  $\varrho$  be the metric of  $X$ . For  $m \in \mathbb{N}$  define (equivalent) Bowen's metric  $\varrho_m^f$  on  $X$  as in Introduction. We write  $B_m^f(x, \varepsilon)$ ,  $r_m^f(\varepsilon, K)$ , and  $s_m^f(\varepsilon, K)$  instead of  $B_{\varrho_m^f}(x, \varepsilon)$ ,  $r_{\varrho_m^f}(\varepsilon, X)$ , and  $s_{\varrho_m^f}(\varepsilon, X)$ . A subsets  $A$  of  $X$  is called  $(m, \varepsilon)$ -spanning or  $(m, \varepsilon)$ -separated if it is  $\varepsilon$ -spanning or  $\varepsilon$ -separated with respect to  $\varrho_m^f$ . By Bowen's definition of the *topological entropy*,

$$h_{\text{top}}(f) = \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} (1/n) \log r_n(\varepsilon, X) = \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} (1/n) \log s_n(\varepsilon, X).$$

**2.1. Local correlation entropy.** Let  $X = (X, \varrho)$  be a compact metric space with a metric  $\varrho$ , and let  $f : X \rightarrow X$  be a continuous map. For  $m \in \mathbb{N}$ ,  $x \in X$ ,  $\varepsilon > 0$ , and  $n \in \mathbb{N}$  define the *correlation sum*  $C_m^f(x, n, \varepsilon)$  by

$$C_m^f(x, n, \varepsilon) = \frac{1}{n^2} \text{card} \{ (i, j) : 0 \leq i, j < n, \varrho_m^f(f^i(x), f^j(x)) \leq \varepsilon \}.$$

Recall the definition (1.5) of the *upper* and *lower local correlation entropies*  $\bar{h}_{\text{cor}}(f, x)$  and  $h_{\text{cor}}(f, x)$  of  $f$  at  $x$ . If  $\bar{h}_{\text{cor}}(f, x) = h_{\text{cor}}(f, x)$  then we say that the local correlation entropy  $h_{\text{cor}}(f, x)$  of  $f$  at  $x$  exists and we put  $h_{\text{cor}}(f, x) = \bar{h}_{\text{cor}}(f, x) = h_{\text{cor}}(f, x)$ . If  $\mu$  is an  $f$ -invariant probability, the *upper* and *lower (measure-theoretic) correlation entropies* (of order 2) of  $f$  with respect to  $\mu$  are defined by (1.4), see e.g. [2, p. 361]. Notice that in this paper we deal solely with correlation entropies of order  $q = 2$ ; for the definition and properties of (measure-theoretic) correlation entropies of arbitrary order  $q$  see e.g. [25, 26, 2].

In the following we summarize some of the known results which will be used later. The first one was in fact proved in [24, p. 355], see also [26, Lemma 2.14].

**Proposition 1** ([24]). *Let  $(X, f)$  be a dynamical system. Then, for every  $f$ -invariant measure  $\mu$ ,*

$$\bar{h}_{\text{cor}}(f, \mu) \leq h_{\mu}(f).$$

Correlation entropy  $\bar{h}_{\text{cor}}(f, \mu)$  can be strictly smaller than measure-theoretic entropy. For example, in [26, Example 2.28] the author constructs a subshift  $(X, \sigma)$  with invariant measure  $\mu$  such that  $h_{\text{cor}}(f, \mu) = 0$  and  $h_{\mu}(f) > 0$ .

The following result was first proved by Pesin [19], see also [20, 1, 22, 15]. (There, the space  $X$  can be any complete separable metric space.)

**Proposition 2** ([19]). *Let  $(X, f)$  be a dynamical system. Then, for every  $f$ -ergodic measure  $\mu$ ,*

$$\mathcal{C}_m^f(x, \varepsilon) = \bar{\mathcal{C}}_m^f(x, \varepsilon) = c_m^f(\mu, \varepsilon)$$

for  $\mu$ -a.e.  $x \in X$  and every  $\varepsilon > 0$  which is a continuity point of  $c_m^f(\mu, \cdot)$ .

As a consequence of Proposition 2 we obtain that, for ergodic  $\mu$ ,

$$(2.1) \quad \bar{h}_{\text{cor}}(f, x) = \bar{h}_{\text{cor}}(f, \mu) \quad \text{and} \quad h_{\text{cor}}(f, x) = h_{\text{cor}}(f, \mu) \quad \text{for } \mu\text{-a.e. } x \in X.$$

For uniquely ergodic systems one can strengthen the previous theorem and obtain convergence of correlation sums to correlation integral for every point.

**Proposition 3.** *Let  $(X, f)$  be a uniquely ergodic dynamical system and  $\mu$  be the unique  $f$ -invariant measure. Then*

$$\mathcal{C}_m^f(x, \varepsilon) = \bar{\mathcal{C}}_m^f(x, \varepsilon) = c_m^f(\mu, \varepsilon)$$

for every  $x \in X$  and every  $\varepsilon > 0$  which is a continuity point of  $c_m^f(\mu, \cdot)$ .

*Proof.* One can repeat e.g. the proof from [15], with applying the Birkhoff ergodic theorem (to define  $Y \subseteq X$  of full measure) replaced by applying Oxtoby's theorem [17, (5.3)] (see also [7, Theorem 4.10]) and putting  $Y = X$ .  $\square$

**2.2. Correlation dimension.** Correlation dimension [9, 10] is another widely used characteristic based on the correlation integral. Recall that upper and lower correlation dimensions (of order 2) are defined by

$$(2.2) \quad \bar{d}_{\text{cor}}(f, \mu) = \limsup_{\varepsilon \rightarrow 0} \frac{\log \mathcal{C}_1^f(\mu, \varepsilon)}{-\log \varepsilon}, \quad d_{\text{cor}}(f, \mu) = \liminf_{\varepsilon \rightarrow 0} \frac{\log \mathcal{C}_1^f(\mu, \varepsilon)}{-\log \varepsilon}.$$

One can analogously define upper and lower *local correlation dimensions*  $\bar{d}_{\text{cor}}(f, x)$  and  $d_{\text{cor}}(f, x)$  by

$$(2.3) \quad \bar{d}_{\text{cor}}(f, x) = \limsup_{\varepsilon \rightarrow 0} \frac{\log \mathcal{C}_1^f(x, \varepsilon)}{-\log \varepsilon}, \quad d_{\text{cor}}(f, x) = \liminf_{\varepsilon \rightarrow 0} \frac{\log \mathcal{C}_1^f(x, \varepsilon)}{-\log \varepsilon}.$$

**2.3. Shifts and subshifts.** Let  $p \geq 2$  be an integer and let  $\mathcal{A}_p = \{0, 1, \dots, p-1\}$ . Put

$$\Sigma_p = \mathcal{A}_p^{\mathbb{N}_0} = \{x = (x_i)_{i=0}^\infty : x_i \in \mathcal{A}_p \text{ for every } i\}.$$

Define a metric  $\varrho$  on  $\Sigma_p$  by

$$\varrho(x, y) = 2^{-k}, \quad k = \min\{i \geq 0 : x_i \neq y_i\}$$

for  $x \neq y$ , and  $\varrho(x, y) = 0$  for  $x = y$ ; thus  $\varrho(x, y) \leq 1/2$  if and only if  $x_0 = y_0$ . Then  $(\Sigma_p, \varrho)$  is a compact metric space homeomorphic to the Cantor ternary set. The *shift*  $\sigma : \Sigma_p \rightarrow \Sigma_p$  is defined by

$$\sigma((x_i)_i) = (y_i)_i, \quad \text{where } y_i = x_{i+1} \text{ for every } i.$$

The dynamical system  $(\Sigma_p, \sigma)$  is called the *(one-sided) full shift* on  $p$  symbols. If  $X \subseteq \Sigma_p$  is a nonempty closed  $\sigma$ -invariant set then the restriction  $\sigma|_X : X \rightarrow X$  is called a *subshift*; since no confusion can arise, the restriction  $\sigma|_X$  will be denoted by  $\sigma$ .

The members of  $\mathcal{A}_p^* = \bigcup_{k \geq 0} \mathcal{A}_p^k$  are called *words*. Let  $k \geq 0$  and  $w = w_0 \dots w_{k-1} \in \mathcal{A}_p^k$ . Then we say that  $w$  is a *k-word* and that the *length* of it is  $|w| = k$ . The *cylinder*  $[w]$  is the clopen set  $\{x \in \Sigma_p : x_i = w_i \text{ for every } 0 \leq i < k\}$ .

For a  $\sigma$ -invariant measure  $\mu$  put

$$(2.4) \quad \tilde{\mu}(k) = \sum_{w \in \mathcal{A}_p^k} (\mu([w]))^2.$$

The next two lemmas (for the second one see e.g. [25, p. 774]) follows from the fact that  $\varrho_m^\sigma(y, z) \leq 2^{-k}$  if and only if  $\varrho(y, z) \leq 2^{-(k+m-1)}$  if and only if there is  $w \in \mathcal{A}_p^{k+m-1}$  such that  $y, z \in [w]$ .

**Lemma 4.** *Let  $(X, \sigma)$  be a subshift and  $\varepsilon \in (0, 1]$ . Let  $k \geq 0$  be an integer such that  $\varepsilon \in [2^{-k}, 2^{-(k-1)})$ . Then, for every  $x \in X$  and  $m, n \in \mathbb{N}$ ,*

$$C_m^\sigma(x, n, \varepsilon) = C_1^\sigma(x, n, 2^{-(k+m-1)}).$$

Consequently,

$$\bar{c}_m^\sigma(x, \varepsilon) = \bar{c}_1^\sigma(x, 2^{-(k+m-1)}) \quad \text{and} \quad \underline{c}_m^\sigma(x, \varepsilon) = \underline{c}_1^\sigma(x, 2^{-(k+m-1)}).$$

**Lemma 5.** *Let  $(X, \sigma)$  be a subshift,  $\mu$  be a  $\sigma$ -invariant measure, and  $\varepsilon \in (0, 1]$ . Let  $k \geq 0$  be an integer such that  $\varepsilon \in [2^{-k}, 2^{-(k-1)})$ . Then, for every  $m \in \mathbb{N}$ ,*

$$c_m^\sigma(\mu, \varepsilon) = c_1^\sigma(\mu, 2^{-(k+m-1)}) = \tilde{\mu}(k+m-1),$$

and so

$$\bar{h}_{\text{cor}}(\sigma, \mu) = \limsup_{m \rightarrow \infty} (-1/m) \log \tilde{\mu}(m) \quad \text{and} \quad \underline{h}_{\text{cor}}(\sigma, \mu) = \liminf_{m \rightarrow \infty} (-1/m) \log \tilde{\mu}(m).$$

If  $\pi = (\pi_0, \dots, \pi_{p-1})$  is a probability vector (that is,  $\pi_i \geq 0$  and  $\sum_i \pi_i = 1$ ), then the  $(\sigma$ -invariant Borel probability) measure  $\mu$  on  $(\Sigma_p, \mathcal{B}(\Sigma_p))$  such that  $\mu([w]) = \prod_{i < k} \pi_{w_i}$  for every  $k \geq 1$  and  $w \in \mathcal{A}_p^k$ , is called the *Bernoulli measure* generated by  $\pi$ . An easy consequence of Lemma 5 is the following result, see [25, p. 773], [26, Sect. 2.5.2].

**Lemma 6.** *Let  $(\Sigma_p, \sigma)$  be the full shift,  $\pi = (\pi_0, \dots, \pi_{p-1})$  be a probability vector, and  $\mu$  be the Bernoulli measure generated by  $\pi$ . Then*

$$h_{\text{cor}}(\sigma, \mu) = -\log \left( \sum_{i < p} \pi_i^2 \right).$$

**Corollary 7.** *Let  $p \geq 2$  and let  $(\Sigma_p, \sigma)$  be the full shift. Then for every  $h \in [0, \log p]$  there is a Cantor subset  $X_h$  of  $\Sigma_p$  such that*

$$h_{\text{cor}}(\sigma, x) = h \quad \text{for every } x \in X_h.$$

*Proof.* Since  $h \in [0, \log p]$ , there is a probability vector  $\pi = (\pi_0, \dots, \pi_{p-1})$  such that  $\sum_i \pi_i^2 = e^{-h}$ . Let  $\mu$  be the Bernoulli measure generated by  $\pi$ ; note that  $\mu$  is  $\sigma$ -ergodic. By (2.1) and Lemma 6, there is a Borel subset  $Y_h$  of  $\Sigma_p$  such that  $\mu(Y_h) = 1$  and  $h_{\text{cor}}(\sigma, x) = h$  for every  $x \in Y_h$ . Since  $\mu$  is non-atomic,  $Y_h$  is uncountable and hence it contains a Cantor set (see e.g. [23, Theorem 3.2.7]).  $\square$

## 3. PROOF OF THEOREM A

**Lemma 8.** *Let  $X$  be a compact metric space and  $\varepsilon > 0$ . Put  $\eta = 1/r(\varepsilon/2, X)$ . Then for every continuous map  $f : X \rightarrow X$ ,  $x \in X$ , and  $m, n \in \mathbb{N}$ ,*

$$C_m^f(x, n, \varepsilon) \geq \eta^m.$$

*Consequently,  $\bar{c}_m^f(x, \varepsilon) \geq \underline{c}_m^f(x, \varepsilon) \geq \eta^m$  and*

$$\bar{d}_{\text{cor}}(f, x) \leq \bar{d}_{\text{box}}(X), \quad \underline{d}_{\text{cor}}(f, x) \leq \underline{d}_{\text{box}}(X).$$

*Proof.* Put  $p = r(\varepsilon/2, X)$ ,  $\eta = 1/p$ , and take a finite subset  $\{y_0, \dots, y_{p-1}\}$  of  $X$  which  $(\varepsilon/2)$ -spans  $X$ . Fix arbitrary continuous  $f : X \rightarrow X$ ,  $x \in X$ , and  $m, n \in \mathbb{N}$ ; for  $i \geq 0$  denote  $f^i(x)$  by  $x_i$ .

Recall that  $\mathcal{A}_p^m$  is the set of  $m$ -words  $w = w_0 \dots w_{m-1}$  over  $\mathcal{A}_p = \{0, \dots, p-1\}$ . Take a partition  $(N_w)_{w \in \mathcal{A}_p^m}$  of  $\{0, 1, \dots, n-1\}$  such that, for every  $w = w_0 \dots w_{m-1}$ ,

$$N_w \subseteq \{0 \leq i < n-1 : x_{i+h} \in B(y_{w_h}, \varepsilon/2) \text{ for every } 0 \leq h < m\}.$$

Notice that  $\varrho_m^f(x_i, x_j) \leq \varepsilon$  for every  $i, j \in N_w$ . Put  $n_w = |N_w|$ . Since  $\sum_w n_w = n$ , the arithmetic-quadratic mean inequality yields

$$C_m^f(x, n, \varepsilon) \geq \frac{1}{n^2} \cdot \sum_{w \in \mathcal{A}_p^m} n_w^2 \geq \frac{1}{n^2} \cdot \frac{n^2}{p^m} = \eta^m.$$

□

The easy proof of the following lemma is skipped.

**Lemma 9.** *Let  $(X, f)$  be a dynamical system,  $x \in X$ , and  $m \in \mathbb{N}$ . Then*

- (a)  $\bar{c}_m^f(x, \varepsilon)$  and  $\underline{c}_m^f(x, \varepsilon)$  are non-decreasing functions of  $\varepsilon$  and non-increasing functions of  $m$ ;
- (b)  $0 < \underline{c}_m^f(x, \varepsilon) \leq \bar{c}_m^f(x, \varepsilon) \leq 1$  for every  $\varepsilon > 0$ ;
- (c)  $\underline{c}_m^f(x, \varepsilon) = \bar{c}_m^f(x, \varepsilon) = 1$  for every  $\varepsilon \geq \text{diam}_\varrho(X)$ .

The next lemma states that in the limits from (1.5) and (1.6) one can use any sublacunary sequences  $(n_j)_{j \geq 1}$  and  $(m_j)_{j \geq 1}$  of integers.

**Lemma 10.** *Let  $(X, f)$  be a dynamical system,  $m \in \mathbb{N}$ ,  $\varepsilon > 0$ , and  $x \in X$ . Let  $(n_j)_j, (m_j)_j$  be increasing sequences of integers such that  $n_{j+1}/n_j \rightarrow 1$  and  $m_{j+1}/m_j \rightarrow 1$  for  $j \rightarrow \infty$ . Then*

$$\bar{c}_m^f(x, \varepsilon) = \limsup_{j \rightarrow \infty} C_m^f(x, n_j, \varepsilon), \quad \underline{c}_m^f(x, \varepsilon) = \liminf_{j \rightarrow \infty} C_m^f(x, n_j, \varepsilon),$$

and

$$\bar{h}_{\text{cor}}(f, x) = \lim_{\varepsilon \rightarrow 0} \limsup_{j \rightarrow \infty} (-1/m_j) \log \underline{c}_{m_j}^f(x, \varepsilon), \quad h_{\text{cor}}(f, x) = \lim_{\varepsilon \rightarrow 0} \liminf_{j \rightarrow \infty} (-1/m_j) \log \bar{c}_{m_j}^f(x, \varepsilon).$$

*Proof.* If  $n_j \leq n < n_{j+1}$  then

$$\left(\frac{n_j}{n}\right)^2 C_m^f(x, n_j, \varepsilon) \leq C_m^f(x, n, \varepsilon) \leq \left(\frac{n_j}{n}\right)^2 C_m^f(x, n_j, \varepsilon) + \frac{n^2 - n_j^2}{n^2}.$$

Since correlation sums are bounded,  $|C_m^f(x, n, \varepsilon) - C_m^f(x, n_j, \varepsilon)|$  is arbitrarily small for  $j$  large enough. Now the first part of the lemma follows.

For  $m \in \mathbb{N}$  put  $a_m = -\log \underline{c}_m^f(x, \varepsilon)$ . By Lemma 9,  $0 \leq a_m \leq a_{m+1}$  for every  $m$ . Thus

$$\frac{m_j}{m_{j+1}} \cdot \frac{a_{m_j}}{m_j} \leq \frac{a_m}{m} \leq \frac{m_{j+1}}{m_j} \cdot \frac{a_{m_{j+1}}}{m_{j+1}}$$

whenever  $m_j \leq m < m_{j+1}$ . Using this and the fact that  $a_m/m \leq r(\varepsilon/2, X)$  for every  $m$  by Lemma 8, we easily obtain that

$$\limsup_{m \rightarrow \infty} a_m/m = \limsup_{j \rightarrow \infty} a_{m_j}/m_j \quad \text{and} \quad \liminf_{m \rightarrow \infty} a_m/m = \liminf_{j \rightarrow \infty} a_{m_j}/m_j.$$

This proves the second part of the lemma. □

### 3.1. Local correlation entropy of $f^k$ : the lower bound.

**Lemma 11.** *Let  $(X, f)$  be a dynamical system,  $m, h \in \mathbb{N}$ ,  $x \in X$ , and  $\varepsilon > 0$ . Then*

$$\bar{c}_m^f(f^h(x), \varepsilon) = \bar{c}_m^f(x, \varepsilon), \quad \underline{c}_m^f(f^h(x), \varepsilon) = \underline{c}_m^f(x, \varepsilon).$$

*Proof.* For every  $n \in \mathbb{N}$  we easily have

$$(3.1) \quad \left(\frac{n+h}{n}\right)^2 C_m^f(x, n+h, \varepsilon) - \frac{2hn+h^2}{n^2} \leq C_m^f(f^h(x), n, \varepsilon) \leq \left(\frac{n+h}{n}\right)^2 C_m^f(x, n+h, \varepsilon),$$

from which the lemma immediately follows.  $\square$

**Lemma 12.** *Let  $(X, f)$  be a dynamical system and  $k, h \in \mathbb{N}$ . Then for every  $\varepsilon > 0$  there are  $0 < \gamma < \delta < \varepsilon$  such that*

$$(3.2) \quad \bar{c}_m^{f^k}(x, \gamma) \leq \bar{c}_m^{f^k}(f^h(x), \delta) \leq \bar{c}_m^{f^k}(x, \varepsilon), \quad \underline{c}_m^{f^k}(x, \gamma) \leq \underline{c}_m^{f^k}(f^h(x), \delta) \leq \underline{c}_m^{f^k}(x, \varepsilon)$$

for every  $x \in X$  and  $m \in \mathbb{N}$ .

*Proof.* Applying Lemma 11 to  $f^k$  allows us to assume that  $h < k$ . Since  $f^{k-h}$  is uniformly continuous, there is  $\delta \in (0, \varepsilon)$  such that  $\varrho(f^{k-h}(y), f^{k-h}(z)) \leq \varepsilon$  whenever  $\varrho(y, z) \leq \delta$ . This implies that  $\varrho_m^{f^k}(f^{k-h}(y), f^{k-h}(z)) \leq \varepsilon$  for every  $y, z \in X$  with  $\varrho_m^{f^k}(y, z) \leq \delta$ . Thus

$$(3.3) \quad C_m^{f^k}(f^h(x), n, \delta) \leq C_m^{f^k}(f^k(x), n, \varepsilon) \quad \text{for every } n.$$

An analogous application of uniform continuity of  $f^h$  gives that there is  $\gamma \in (0, \delta)$  such that

$$(3.4) \quad C_m^{f^k}(x, n, \gamma) \leq C_m^{f^k}(f^h(x), n, \delta) \quad \text{for every } n.$$

Now (3.3), (3.4), and Lemma 11 yield (3.2).  $\square$

**Corollary 13.** *Let  $(X, f)$  be a dynamical system,  $k, h \in \mathbb{N}$ , and  $x \in X$ . Then*

$$\bar{h}_{\text{cor}}(f^k, f^h(x)) = \bar{h}_{\text{cor}}(f^k, x), \quad h_{\text{cor}}(f^k, f^h(x)) = h_{\text{cor}}(f^k, x).$$

**Lemma 14.** *Let  $(X, f)$  be a dynamical system and  $k \in \mathbb{N}$ . Then for every  $\varepsilon > 0$  there is  $\delta \in (0, \varepsilon)$  such that*

$$\bar{c}_m^f(x, \delta) \leq \bar{c}_{km}^f(x, \varepsilon), \quad \underline{c}_m^f(x, \delta) \leq \underline{c}_{km}^f(x, \varepsilon)$$

for every  $m \in \mathbb{N}$  and  $x \in X$ .

*Proof.* Since  $X$  is compact and  $f$  is continuous, there is  $\delta \in (0, \varepsilon)$  such that  $\varrho(y, z) \leq \delta$  implies  $\varrho(f^h(y), f^h(z)) \leq \varepsilon$  for every  $h = 0, \dots, k-1$ . Hence

$$\varrho_{km}^f(y, z) \leq \varepsilon \quad \text{for every } y, z \in X \text{ with } \varrho_m^{f^k}(y, z) \leq \delta.$$

This gives, for every  $x \in X$  and  $m, n \in \mathbb{N}$ ,

$$C_m^{f^k}(x, n, \delta) \leq C_{km}^f(x, n, \varepsilon).$$

Now the lemma immediately follows.  $\square$

**Corollary 15.** *Let  $(X, f)$  be a dynamical system,  $k \in \mathbb{N}$ , and  $x \in X$ . Then*

$$\bar{h}_{\text{cor}}(f^k, x) \geq k \cdot \bar{h}_{\text{cor}}(f, x), \quad h_{\text{cor}}(f^k, x) \geq k \cdot h_{\text{cor}}(f, x).$$

*Proof.* By Lemmas 12 and 10, for every  $\varepsilon > 0$  there is  $\delta_\varepsilon \in (0, \varepsilon)$  such that

$$\limsup_{m \rightarrow \infty} (-1/m) \log \bar{c}_m^{f^k}(x, \delta_\varepsilon) \geq \limsup_{m \rightarrow \infty} (-1/m) \log \bar{c}_{km}^f(x, \varepsilon) = k \cdot \limsup_{m \rightarrow \infty} (-1/m) \log \bar{c}_m^f(x, \varepsilon).$$

Hence  $\bar{h}_{\text{cor}}(f^k, x) \geq k \cdot \bar{h}_{\text{cor}}(f, x)$ . The second inequality can be proved analogously.  $\square$



**3.2. Local correlation entropy of  $f^k$ : a combinatorial lemma.** Fix a finite set  $V$  consisting of  $n$  points, and a partition  $\mathcal{V} = (V_0, V_1, \dots, V_{k-1})$  of it into  $k \geq 2$  nonempty subsets. Consider an undirected simple (not necessarily connected) graph  $G$  with the set of vertices  $V$ . For  $0 \leq a, b < k$ , an edge  $\{i, j\}$  of  $G$  is called an  $ab$ -edge if  $i \in V_a$  and  $j \in V_b$ , or vice versa. We say that a graph  $G$  is  $\mathcal{V}$ -admissible if the following hold:

$$(3.5) \quad \begin{aligned} &\text{If } \{i, j\}, \{i', j\} \text{ are different edges of } G \text{ with } i, i' \in V_a \text{ and } j \in V_b \text{ (} a \neq b \text{),} \\ &\text{then } \{i, i'\} \text{ is also an edge of } G. \end{aligned}$$

The number of all  $ab$ -edges of  $G$  is denoted by  $m_{ab}(G)$ . Put

$$\kappa(G) = \sum_{a < b} m_{ab}(G) - \sum_a m_{aa}(G).$$

Our aim is to find an upper bound for  $\kappa(G)$  depending only on  $n$  and  $k$ . To this end, we say that a  $\mathcal{V}$ -admissible graph  $G$  is  $\mathcal{V}$ -optimal if  $\kappa(G') \leq \kappa(G)$  for every  $\mathcal{V}$ -admissible graph  $G'$ . Further, if  $G$  is  $\mathcal{V}$ -optimal and the number of edges of every  $\mathcal{V}$ -optimal graph  $G'$  is larger than or equal to that of  $G$ , we say that  $G$  is a *minimal  $\mathcal{V}$ -optimal* graph. The following lemma gives a characterization of minimal  $\mathcal{V}$ -optimal graphs.

**Lemma 16.** *Let  $G$  be a graph with the set of vertices  $V$ . Then  $G$  is a minimal  $\mathcal{V}$ -optimal graph if and only if the following two conditions hold for every  $0 \leq a < b < k$ :*

- (a)  $m_{aa}(G) = 0$  and  $m_{ab}(G) = \min\{|V_a|, |V_b|\}$ ;
- (b) *no two  $ab$ -edges have a common vertex.*

Consequently,

$$\max\{\kappa(G) : G \text{ is } \mathcal{V}\text{-admissible}\} = \sum_{a < b} \min\{|V_a|, |V_b|\}.$$

*Proof.* Fix a minimal  $\mathcal{V}$ -optimal graph  $G$  and any  $a \neq b$  with  $|V_a| \leq |V_b|$ . For  $i \in V_a$  define

$$\begin{aligned} A_{ib} &= \{j \in V_b : \{i, j\} \text{ is an edge of } G\}, \\ B_{ib} &= \{i' \in V_a : \{i', j\} \text{ is an edge of } G \text{ for some } j \in A_{ib}\}. \end{aligned}$$

Assume that  $A_{ib} \neq \emptyset$ . Take the ( $\mathcal{V}$ -admissible) graph  $\tilde{G}$  created from  $G$  by removing all  $ab$ -edges  $\{i, j\}$  (with  $j \in A_{ib}$ ) as well as all  $aa$ -edges  $\{i, i'\}$  (with  $i' \in B_{ib} \setminus \{i\}$ ). Then  $\kappa(\tilde{G}) = \kappa(G) - |A_{ib}| + (|B_{ib}| - 1)$  since  $i \in B_{ib}$ . By minimality of  $G$  we have that  $\kappa(\tilde{G}) < \kappa(G)$  and so  $|A_{ib}| \geq |B_{ib}|$ . If  $A_{ib} = \emptyset$  then  $B_{ib} = \emptyset$  by the definition of  $B_{ib}$ . Thus, in both cases,

$$(3.6) \quad |A_{ib}| \geq |B_{ib}|.$$

Assume again that  $A_{ib} \neq \emptyset$ . Take any  $j \in A_{ib}$  and define  $A_{ja}, B_{ja}$  analogously. Then  $B_{ja} \supseteq A_{ib}$  and  $B_{ib} \supseteq A_{ja}$ . Inequality (3.6), applied also to  $j$  and  $a$ , yields  $|A_{ib}| \leq |B_{ja}| \leq |A_{ja}| \leq |B_{ib}| \leq |A_{ib}|$ . Thus

$$|A_{ib}| = |B_{ib}|$$

and, for every  $j \in A_{ib}$ ,

$$A_{ja} = B_{ib}, \quad B_{ja} = A_{ib}.$$

$\mathcal{V}$ -admissibility of  $G$  now gives that  $A_{ib} \cup B_{ib}$  is a clique of  $G$  (that is, the induced subgraph is complete). Since  $G$  is minimal, this easily implies that  $A_{ib}$  is a singleton. (For if not, there is  $l \geq 2$  such that we can write  $A_{ib} = \{i_1 = i, i_2, \dots, i_l\}$  and  $B_{ib} = \{j_1 = j, j_2, \dots, j_l\}$ . Create a graph  $\tilde{G}$  from  $G$  by removing  $l(l-1)$  edges  $\{i_r, i_s\}$ ,  $\{j_r, j_s\}$  ( $r \neq s$ ) and  $l(l-1)$  edges  $\{i_r, j_s\}$  ( $r \neq s$ ). Then  $\tilde{G}$  is  $\mathcal{V}$ -admissible,  $\kappa(\tilde{G}) = \kappa(G)$ , and  $\tilde{G}$  has smaller number of edges than  $G$  — a contradiction.)

We have proved that, for every  $a \neq b$  with  $|V_a| \leq |V_b|$  and for every  $i \in V_a$ , there is either zero or exactly one  $ab$ -edge from  $i$ . Suppose that there is  $i \in V_a$  with no  $ab$ -edge from it; since  $|V_a| \leq |V_b|$ , there is also  $j \in V_b$  not adjacent to any vertex from  $V_a$ . But then the  $\mathcal{V}$ -admissible graph  $\tilde{G}$  obtained from  $G$  by adding the edge  $\{i, j\}$  has  $\kappa(\tilde{G}) = \kappa(G) + 1$ , a contradiction. Hence, for every  $i \in V_a$  there is exactly one  $j \in V_b$  adjacent to it. Thus  $m_{ab}(G) = |V_a| = \min\{|V_a|, |V_b|\}$ .

If there are  $i \neq i'$  from  $V_a$  adjacent to the same  $j \in V_b$ , then (3.5) implies that  $\{i, i'\}$  is an edge of  $G$ . But then the  $\mathcal{V}$ -admissible graph  $\tilde{G}$ , created from  $G$  by removing edges  $\{i, i'\}$  and  $\{i', j\}$ , has  $\kappa(\tilde{G}) = \kappa(G)$ ,



which contradicts minimality of  $G$ . So we have (b). Finally,  $\mathcal{V}$ -optimality of  $G$  now gives that  $m_{aa}(G) = 0$ . Hence (a) is proved too.

Now take any graph  $G$  with the set of vertices  $V$ , which satisfies (a) and (b). By (b),  $G$  is  $\mathcal{V}$ -admissible (indeed, the condition (3.5) is trivially satisfied). By (a),  $\kappa(G) = \sum_{a < b} \min\{|V_a|, |V_b|\}$ . Thus, by the previous part of the proof,  $G$  is  $\mathcal{V}$ -optimal. Further, again by the previous part of the proof, every minimal  $\mathcal{V}$ -optimal graph  $G'$  has the number of edges equal to that of  $G$ . Thus  $G$  is a minimal  $\mathcal{V}$ -optimal graph.  $\square$

**Lemma 17.** *Let  $V$  be a finite set of cardinality  $n$  and  $\mathcal{V}$  be a partition of it into  $k \geq 2$  nonempty subsets. Then*

$$\kappa(G) \leq \frac{n(k-1)}{2}$$

for every  $\mathcal{V}$ -admissible graph  $G$ .

*Proof.* We first prove that

$$(3.7) \quad \sum_{h=0}^{k-1} hx_h \leq \frac{k-1}{2}$$

for every  $x \in K = \{(x_0, \dots, x_{k-1}) \in \mathbb{R}^k : \sum_h x_h = 1, x_0 \geq \dots \geq x_{k-1} \geq 0\}$ . To this end, define a map  $f : K \rightarrow \mathbb{R}$  by  $f(x) = \sum_{h=0}^{k-1} hx_h$ . Since  $K$  is compact and  $f$  is continuous, there is  $\bar{x} \in K$  which maximizes  $f$ . Suppose that  $\bar{x}_h > \bar{x}_{h+1}$  for some  $h < k-1$ . Define  $x' \in \mathbb{R}^k$  by  $x'_i = (\bar{x}_h + \bar{x}_{h+1})/2$  if  $i \in \{h, h+1\}$ , and  $x'_i = \bar{x}_i$  otherwise. Then  $x' \in K$  and  $f(x') = f(\bar{x}) + (\bar{x}_h - \bar{x}_{h+1})/2 > f(\bar{x})$ , a contradiction. Thus  $\bar{x}_h = (1/k)$  for every  $h$  and (3.7) follows

Now we can prove Lemma 17. Put  $n_h = |V_h|$  for  $h = 0, \dots, k-1$ ; we may assume that  $n_0 \geq n_1 \geq \dots \geq n_{k-1}$ . Let  $G$  be a  $\mathcal{V}$ -admissible graph. By Lemma 16 and (3.7) with  $x_h = n_h/n$ ,

$$\kappa(G) \leq \sum_{h=0}^{k-1} hn_h = n \cdot \sum_{h=0}^{k-1} hx_h \leq \frac{n(k-1)}{2}.$$

$\square$

### 3.3. Local correlation entropy of $f^k$ : the upper bound.

**Lemma 18.** *Let  $(X, f)$  be a dynamical system,  $k \geq 2$ ,  $\varepsilon > 0$ ,  $x \in X$ , and  $m, n \in \mathbb{N}$ . Then*

$$C_{km}^f(x, kn, \varepsilon) \leq \frac{k-2}{kn} + \frac{2}{k^2} \sum_{h=0}^{k-1} C_m^{f^k}(f^h(x), n, 2\varepsilon).$$

*Proof.* Put  $\bar{n} = kn$ ,  $V = \{0, 1, \dots, \bar{n}-1\}$  and, for  $0 \leq a < k$ ,  $V_a = \{i \in V : i \equiv a \pmod{k}\}$ . Let  $G$  be an undirected simple graph with the set of vertices  $V$  and such that, for any  $i \neq j$  from  $V$ ,  $\{i, j\}$  is an edge of  $G$  if and only if

$$\varrho_{km}^f(f^i(x), f^j(x)) \leq \begin{cases} 2\varepsilon & \text{if } i, j \in V_a \text{ for some } a; \\ \varepsilon & \text{otherwise.} \end{cases}$$

Notice that the number  $m(G)$  of edges of  $G$  satisfies

$$(3.8) \quad m(G) \geq \frac{1}{2} \left[ \bar{n}^2 C_{km}^f(x, \bar{n}, \varepsilon) - \bar{n} \right].$$

Further,  $G$  is  $\mathcal{V}$ -admissible. In fact, fix any  $a \neq b$ , different  $i, i' \in V_a$ , and  $j \in V_b$ . If  $\{i, j\}$ ,  $\{i', j\}$  are  $ab$ -edges, then  $\varrho_{km}^f(f^i(x), f^j(x)) \leq \varepsilon$  and  $\varrho_{km}^f(f^{i'}(x), f^j(x)) \leq \varepsilon$ . Hence, by the triangle inequality,  $\varrho_{km}^f(f^i(x), f^{i'}(x)) \leq 2\varepsilon$  and so  $\{i, i'\}$  is an edge of  $G$ . Lemma 17 yields

$$(3.9) \quad m(G) = \sum_a m_{aa}(G) + \sum_{a < b} m_{ab}(G) \leq 2 \sum_a m_{aa}(G) + \frac{\bar{n}(k-1)}{2}.$$

Since  $\varrho_m^{f^k} \leq \varrho_{km}^f$ , for every  $0 \leq a < k$  the definition of  $G$  gives

$$n^2 \cdot C_m^{f^k}(f^a(x), n, 2\varepsilon) \geq 2m_{aa}(G) + n.$$

This together with (3.8) and (3.9) yield

$$\frac{1}{2} \left[ \bar{n}^2 C_{km}^f(x, \bar{n}, \varepsilon) - \bar{n} \right] \leq m(G) \leq \sum_a \left[ n^2 \cdot C_m^{f^k}(f^a(x), n, 2\varepsilon) - n \right] + \frac{\bar{n}(k-1)}{2}.$$

Now a simple computation gives the desired inequality.  $\square$

**Lemma 19.** *Let  $(X, f)$  be a dynamical system and  $0 \leq h < k$  be integers. Then for every  $\varepsilon > 0$  there is  $\eta(\varepsilon) > 0$  such that*

$$(3.10) \quad \lim_{\varepsilon \rightarrow 0} \eta(\varepsilon) = 0 \quad \text{and} \quad C_m^{f^k}(f^h(x), n, \varepsilon) \leq C_m^{f^k}(x, n+1, \eta(\varepsilon)) + \frac{3}{n}$$

for every  $x \in X$  and  $m, n \in \mathbb{N}$ .

*Proof.* In the proof of Lemma 12 we have shown that for every  $e > 0$  there is  $d(e) \in (0, e)$  such that  $C_m^{f^k}(f^h(x), n, d(e)) \leq C_m^{f^k}(f^k(x), n, e)$ ; see (3.3). Fix a sequence  $(e_i)_i$  decreasing to zero and put  $d_i = d(e_i)$ ; we may assume that  $d_i > d_{i+1}$  for every  $i$ . For every  $\varepsilon > 0$  define

$$\eta(\varepsilon) = \begin{cases} e_i & \text{if } \varepsilon \in (d_{i+1}, d_i] \text{ for some } i; \\ \text{diam}(X) & \text{if } \varepsilon > d_0. \end{cases}$$

Then, by Lemma 9,  $C_m^{f^k}(f^h(x), n, \varepsilon) \leq C_m^{f^k}(f^k(x), n, \eta(\varepsilon))$  for every  $\varepsilon > 0$ . Combining this with (3.1), applied to  $f' = f^k$  and  $h' = 1$ , yields

$$C_m^{f^k}(f^h(x), n, \varepsilon) \leq \left( \frac{n+1}{n} \right)^2 C_m^{f^k}(x, n+1, \eta(\varepsilon)) \leq C_m^{f^k}(x, n+1, \eta(\varepsilon)) + \frac{3}{n}.$$

Since  $\lim_{\varepsilon} \eta(\varepsilon) = 0$  is immediate by the choice of  $\eta$ , the lemma is proved.  $\square$

*Proof of Theorem A.* We may assume that  $k \geq 2$ . Lemma 19, applied to every  $h \in \{0, \dots, k-1\}$ , gives that for every  $\varepsilon > 0$  there is  $\eta(\varepsilon) > 0$  such that  $\lim_{\varepsilon \rightarrow 0} \eta(\varepsilon) = 0$  and

$$C_m^{f^k}(f^h(x), n, \varepsilon) \leq C_m^{f^k}(x, n+1, \eta(\varepsilon)) + \frac{3}{n}$$

for every  $0 \leq h < k$ ,  $x \in X$ , and  $m, n \in \mathbb{N}$ . Now, by Lemma 18,

$$C_{km}^f(x, kn, \varepsilon) \leq \frac{k-2}{kn} + \frac{2}{k^2} \sum_{h=0}^{k-1} C_m^{f^k}(f^h(x), n, 2\varepsilon) \leq \frac{k+4}{kn} + \frac{2}{k} C_m^{f^k}(x, n+1, \eta(2\varepsilon)).$$

By taking the limit as  $n$  approaches infinity, and using Lemma 10 we obtain

$$\bar{c}_{km}^f(x, \varepsilon) \leq \frac{2}{k} \cdot \bar{c}_m^{f^k}(x, \eta(2\varepsilon)), \quad \underline{c}_{km}^f(x, \varepsilon) \leq \frac{2}{k} \cdot \underline{c}_m^{f^k}(x, \eta(2\varepsilon)).$$

Consequently, again using Lemma 10,

$$k \cdot \bar{h}_{\text{cor}}(f, x) \geq \bar{h}_{\text{cor}}(f^k, x), \quad k \cdot \underline{h}_{\text{cor}}(f, x) \geq \underline{h}_{\text{cor}}(f^k, x).$$

Since the opposite inequalities were shown in Corollary 15, Theorem A is proved.  $\square$

#### 4. PROOF OF THEOREM B

**Lemma 20.** *Let  $(X, f)$  be a dynamical system,  $x \in X$ ,  $\varepsilon > 0$ , and  $m, n \in \mathbb{N}$ . Then*

$$C_m^f(x, n, \varepsilon) \geq \frac{1}{r_m(\varepsilon/2, X)}.$$

*Proof.* The proof is pretty similar to that of Lemma 8; the only difference is that instead of  $(\varepsilon/2)$ -spanning sets we use  $(m, \varepsilon/2)$ -spanning sets. For completeness, the details follow.

Let  $\{y_0, \dots, y_{p-1}\}$  be an  $(m, \varepsilon/2)$ -spanning subset of minimal cardinality  $p = r_m(\varepsilon/2, X)$ . Hence for every  $i \geq 0$  and  $x_i = f^i(x)$  there is  $v_i$  with  $\varrho_m^f(x_i, y_{v_i}) \leq \varepsilon/2$ . For  $0 \leq v < p$  put

$$N_v = \{0 \leq i \leq n-m : v_i = v\} \quad \text{and} \quad n_v = |N_v|.$$

Then, by the arithmetic-quadratic mean inequality,

$$C_m^f(x, n, \varepsilon) \geq \frac{1}{n^2} \cdot \sum_{v < p} n_v^2 \geq \frac{1}{n^2} \cdot \frac{n^2}{p} = \frac{1}{r_m(\varepsilon/2, X)}.$$

□

**Proposition 21.** *Let  $(X, f)$  be a dynamical system and  $x \in X$ . Then*

$$\underline{h}_{\text{cor}}(f, x) \leq \bar{h}_{\text{cor}}(f, x) \leq h_{\text{top}}\left(f|_{\overline{\text{Orb}}_f(x)}\right) \leq h_{\text{top}}(f).$$

The part corresponding to the lower local correlation entropy was proved in [24, p. 354]. The proof used the fact that if  $x$  is a quasi-generic point [4, (4.4)] of an invariant measure  $\mu$ , then [24, p. 355]

$$\underline{h}_{\text{cor}}(f, x) \leq \underline{h}_{\text{cor}}(f, \mu) \leq h_\mu(f).$$

*Proof.* By Lemma 20 and Bowen's definition of topological entropy,

$$\bar{h}_{\text{cor}}(f, x) = \lim_{\varepsilon \rightarrow 0} \limsup_{m \rightarrow \infty} (-1/m) \log \mathcal{C}_m^f(x, \varepsilon) \leq \lim_{\varepsilon \rightarrow 0} \limsup_{m \rightarrow \infty} (1/m) \log r_m(\varepsilon/2, X) = h_{\text{top}}(f).$$

Applying this to  $X' = \overline{\text{Orb}}_f(x)$  and  $f' = f|_{X'}$  yields the required inequality. □

*Remark 22.* Proposition 21 is tightly connected with the fact that, for every  $f$ -invariant measure  $\mu$ ,  $\bar{h}_{\text{cor}}(f, \mu) \leq h_\mu(f)$  (see Proposition 1). Thus, by (2.1),

$$\bar{h}_{\text{cor}}(f, x) \leq h_\mu(f) \quad \text{for } \mu\text{-a.e. } x \in X$$

provided  $\mu$  is ergodic.

Now we embark on the proof of the fact that, for dynamical systems on graphs, local correlation entropies can be arbitrarily close to the topological entropy.

**Proposition 23.** *Let  $(X, f)$  be a dynamical system having a subsystem  $(Y, f)$ , which is a topological extension of the full shift  $(\Sigma_p, \sigma)$  for some  $p \geq 2$ . Then there is  $\varepsilon_0 > 0$  such that the following is true: For every  $\alpha \in \Sigma_p$  there is  $x_\alpha \in Y$  such that  $x_\alpha \neq x_\beta$  whenever  $\alpha \neq \beta$ , and*

$$C_m^f(x_\alpha, n, \varepsilon) \leq C_m^\sigma(\alpha, n, 1/2) \quad \text{for every } \varepsilon \in (0, \varepsilon_0] \text{ and } m, n \in \mathbb{N}.$$

Consequently,

$$\bar{h}_{\text{cor}}(f, x_\alpha) \geq \bar{h}_{\text{cor}}(\sigma, \alpha) \quad \text{and} \quad \underline{h}_{\text{cor}}(f, x_\alpha) \geq \underline{h}_{\text{cor}}(\sigma, \alpha).$$

*Proof.* Let  $h : (Y, f) \rightarrow (\Sigma_p, \sigma)$  be a factor map (that is,  $h$  is a continuous surjection and  $h \circ f = \sigma \circ h$ ). For every  $j \in \mathcal{A}_p = \{0, \dots, p-1\}$  put  $Y_j = h^{-1}([j])$  (recall that  $[j]$  denotes the cylinder  $\{\alpha \in \Sigma_p : \alpha_0 = j\}$ ); this is a closed, hence compact set. Put  $\varepsilon_0 = (1/2) \min_{i \neq j} \text{dist}(Y_i, Y_j)$ ; since the sets  $Y_j$  are pairwise disjoint and compact, we have  $\varepsilon_0 > 0$ .

Fix any  $\alpha = \alpha_0 \alpha_1 \dots \in \Sigma_p$  and take arbitrary  $x = x_\alpha \in h^{-1}(\{\alpha\})$ ; clearly,  $x_\alpha \neq x_\beta$  whenever  $\alpha \neq \beta$ . Realize that  $f^i(x) \in Y_{\alpha_i}$  for every  $i$ . Hence, by the choice of  $\varepsilon_0$ ,  $\varrho(f^i(x), f^j(x)) \leq \varepsilon_0$  implies  $\alpha_i = \alpha_j$ . Thus also, for every  $i, j$ ,

$$\varrho_m^f(f^i(x), f^j(x)) \leq \varepsilon_0 \quad \text{implies} \quad \tilde{\varrho}_m^\sigma(\sigma^i(\alpha), \sigma^j(\alpha)) \leq 1/2$$

(where  $\tilde{\varrho}$  denotes the metric on  $\Sigma_p$ , see §2.3; recall that  $\tilde{\varrho}(\alpha, \beta) \leq 1/2$  is equivalent to  $\alpha_0 = \beta_0$ ). Now  $C_m^f(x, n, \varepsilon) \leq C_m^\sigma(\alpha, n, 1/2)$  for every  $\varepsilon \in (0, \varepsilon_0]$  and  $m, n \in \mathbb{N}$ , from which the first assertion immediately follows.

The second assertion then follows by Lemma 4. To see this, assume that  $\varepsilon_0 \leq 1$ . For every  $\varepsilon \in (0, \varepsilon_0]$  denote by  $k_\varepsilon$  the unique nonnegative integer such that  $\varepsilon \in [2^{-k_\varepsilon}, 2^{-(k_\varepsilon-1)})$ . Then, by Lemma 4,  $C_m^\sigma(\alpha, n, 1/2) = C_{m-k_\varepsilon+1}^\sigma(\alpha, n, 2^{-k_\varepsilon}) = C_{m-k_\varepsilon+1}^\sigma(\alpha, n, \varepsilon)$ . So, by the first part of the lemma,

$$\limsup_{m \rightarrow \infty} (-1/m) \log \mathcal{C}_m^f(x_\alpha, \varepsilon) \geq \limsup_{m \rightarrow \infty} (-1/m) \log \mathcal{C}_{m-k_\varepsilon+1}^\sigma(\alpha, \varepsilon) = \limsup_{m \rightarrow \infty} (-1/m) \log \mathcal{C}_m^\sigma(\alpha, \varepsilon)$$

and so  $\bar{h}_{\text{cor}}(f, x_\alpha) \geq \bar{h}_{\text{cor}}(\sigma, \alpha)$ . Analogously for lower entropies. □

Recall that subsets  $X_0, \dots, X_{p-1}$  of  $X$  form a *strict  $p$ -horseshoe* of a dynamical system  $(X, f)$  if the sets  $X_i$  are nonempty, closed, pairwise disjoint, and  $f(X_i) \supseteq \bigcup_j X_j$  for every  $0 \leq i < p$ .

**Lemma 24.** *Let  $(X, f)$  be a dynamical system containing a strict  $p$ -horseshoe  $X_0, \dots, X_{p-1}$  for some  $p \geq 2$ . Then  $(X, f)$  has a subsystem  $(Y, f)$  which is a topological extension of the full shift  $(\Sigma_p, \sigma)$ .*

*Proof.* This is standard: Since the sets  $X_0, \dots, X_{p-1}$  form a strict  $p$ -horseshoe, in a usual way for every  $k \geq 2$  we can construct disjoint nonempty compact subsets  $X_a$  ( $a \in \mathcal{A}_p^k$ ) such that

$$f(X_{a_0 a_1 \dots a_{k-1}}) = X_{a_1 a_2 \dots a_{k-1}} \quad \text{and} \quad X_{a_0 a_1 \dots a_{k-1}} \subseteq X_{a_0 a_1 \dots a_{k-2}}$$

for every  $a = a_0 a_1 \dots a_{k-1} \in \mathcal{A}_p^k$ . For  $\alpha = \alpha_0 \alpha_1 \dots \in \Sigma_p$  put  $X_\alpha = \bigcap_{k \geq 1} X_{\alpha_0 \dots \alpha_{k-1}}$ . Then  $Y = \bigcup_\alpha X_\alpha$  is compact,  $\sigma(Y) = Y$ , and  $(Y, f|_Y)$  is a topological extension of the full shift  $(\Sigma_p, \sigma)$ .  $\square$

Now we are ready to prove Theorem B.

*Proof of Theorem B.* By Proposition 21 it suffices to prove the second part of the theorem. We may assume that  $h_{\text{top}}(f) > 0$ . Take arbitrary  $0 < h < h_{\text{top}}(f)$ . By [14] there are integers  $p, k$  with  $(1/k) \log p \geq h$  such that  $f^k$  has a strict  $p$ -horseshoe. By Corollary 7, Lemma 24, and Proposition 23, there is a Cantor set  $X_h$  such that  $\underline{h}_{\text{cor}}(f^k, x) \geq \log p$  for every  $x \in X_h$ . Hence, by Theorem A,  $\underline{h}_{\text{cor}}(f, x) = (1/k) \underline{h}_{\text{cor}}(f^k, x) \geq (1/k) \log p \geq h$  for every  $x \in X_h$ .  $\square$

*Remark 25* (Infimum of local correlation entropies). Since graph maps with positive entropy have (dense) periodic points, for continuous graph maps  $f : X \rightarrow X$  we always have

$$\inf_{x \in X} \underline{h}_{\text{cor}}(f, x) = \inf_{x \in X} \bar{h}_{\text{cor}}(f, x) = 0.$$

The following two examples show that it can happen that the local correlation entropy at every point is strictly smaller than the topological entropy of  $f$  and that, in positive entropy systems on graphs, the set of those  $x$  with positive local correlation entropy can be negligible from the measure-theoretic point of view.

*Example 26.* Take  $\lambda \in (0, \infty]$ . For  $n = 1, 2, \dots$  let  $I_n = [1/(n+1), 1/n]$  and let  $f_n : I_n \rightarrow I_n$  be such that it fixes the end points of  $I_n$ ,  $h_{\text{top}}(f_n) < \lambda$  and  $\sup_n h_{\text{top}}(f_n) = \lambda$ . Define a map  $f : I \rightarrow I$  by

$$f(0) = 0, \quad f(x) = f_n(x) \text{ if } x \in I_n, n \geq 1.$$

Then  $f$  is continuous and  $h_{\text{top}}(f) = \lambda$  (see e.g. [18, Theorem 11.2]). On the other hand, for every  $x$  we have  $\bar{h}_{\text{cor}}(f, x) < \lambda$ . In fact, if  $x = 0$  then  $h_{\text{cor}}(f, x) = 0$  since  $x$  is fixed, and if  $x \in I_n$  then  $\bar{h}_{\text{cor}}(f, x) \leq h_{\text{top}}(f_n) < \lambda$  by Proposition 21.

*Example 27.* Let  $f : [-1, 1] \rightarrow [-1, 1]$  be defined by  $f(x) = 1 - \alpha x^2$ , where  $\alpha \in (1, 2)$  is such that  $1 - \alpha(1 - \alpha)^2 = 0$ . Then almost every point  $x$  is attracted by the 3-periodic orbit of the point 0 [3, p. 119] and hence  $h_{\text{cor}}(f, x) = 0$ . On the other hand, having a point with period 3, the topological entropy of  $f$  is positive.

## 5. UNIQUELY ERGODIC SYSTEMS

In this section we summarize facts on uniquely ergodic systems, which will be used in Section 6. Following [12], we say that a set  $A \subseteq \mathbb{N}_0$  is *uniform Cesàro with density*  $\alpha \geq 0$  if for every  $\varepsilon > 0$  there is  $n_0 \in \mathbb{N}$  such that, for every  $n \geq n_0$  and  $j \in \mathbb{N}_0$ ,

$$(5.1) \quad \left| \frac{1}{n} \cdot |A \cap [j, j+n)| - \alpha \right| < \varepsilon.$$

In such a case the density  $\alpha$  of  $A$  will be denoted by  $d(A)$ . It is easy to check that  $A \subseteq \mathbb{N}_0$  is uniform Cesàro with density  $\alpha$  if and only if there is  $l \in \mathbb{N}$  such that for every  $\varepsilon > 0$  there is  $n_0 \in \mathbb{N}$  with

$$(5.2) \quad \left| \frac{1}{ln} \cdot |A \cap [lj, lj+ln)| - \alpha \right| < \varepsilon \quad \text{for every } n \geq n_0 \text{ and } j \in \mathbb{N}_0.$$

Let  $p \geq 2$ . For words  $u, v \in \mathcal{A}_p^*$  with  $|u| \leq |v|$  and an integer  $l \geq 1$  put

$$N_v^{(l)}(u) = \{i \in \mathbb{N}_0 : v[i, i+l|u|) = u\}, \quad \tau_v^{(l)}(u) = \frac{1}{\lfloor |v|/l \rfloor} \cdot |N_v^{(l)}(u)|;$$

so  $\tau_v^{(l)}(u) \leq 1$  is the frequency of occurrences of  $u$  in  $v$  at positions which are multiples of  $l$ . (Since  $N_v^{(l)}(u) \subseteq [0, \lfloor (|v| - |u|)/l \rfloor]$ , in the definition of  $\tau_v^{(l)}(u)$  we should divide by  $1 + \lfloor (|v| - |u|)/l \rfloor$ ; the difference

is, of course, asymptotically negligible.) If  $u \in \mathcal{A}_p^*$  and  $x \in \Sigma_p$ , define  $N_x^{(l)}(u)$  analogously. For abbreviation, we often write  $N_x, \tau_x$  and  $N_v, \tau_v$  instead of  $N_x^{(1)}, \tau_x^{(1)}$  and  $N_v^{(1)}, \tau_v^{(1)}$ . Note that  $N_x^{(l)}(u)$  is uniform Cesàro if and only if for every  $u \in \mathcal{A}_p^*$  the limit  $\lim_n \tau_{x[jl, (j+n)l]}^{(l)}(u)$  exists uniformly in  $j$  and does not depend on  $j$ ; in such a case we have

$$(5.3) \quad d(N_x^{(l)}(u)) = (1/l) \lim_{n \rightarrow \infty} \tau_{x[jl, (j+n)l]}^{(l)}(u) \quad \text{for every } j.$$

By [13, Theorem 3.9] we have the following.

**Lemma 28** ([13]). *Let  $x \in \Sigma_p$  be almost periodic. Assume that  $N_x(u)$  is uniform Cesàro for every  $u \in \mathcal{A}_p^*$ . Then the subshift  $(\overline{\text{Orb}}_\sigma(x), \sigma)$  is strictly ergodic. Moreover,*

$$\mu([u]) = d(N_x(u)) = \lim_{n \rightarrow \infty} \tau_{x[j, j+n]}(u) \quad \text{for every } u \in \mathcal{A}_p^* \text{ and } j \in \mathbb{N}_0,$$

where  $\mu$  is the unique invariant measure of  $(\overline{\text{Orb}}_\sigma(x), \sigma)$ .

The following lemma gives a condition on  $x$  implying strict ergodicity.

**Lemma 29.** *Let  $x \in \Sigma_p$  be almost periodic and let  $(l_j)_{j \geq 1}$  be an increasing sequence of positive integers with every  $l_{j+1}$  being a multiple of  $l_j$ . Assume that, for every  $j \geq 1$  and every  $l_j$ -word  $v$ , the set*

$$N_x^{(l_j)}(v) = \{i \in \mathbb{N}_0 : x[il_j, (i+1)l_j] = v\}$$

*is uniform Cesàro. Then the subshift  $(\overline{\text{Orb}}_\sigma(x), \sigma)$  is strictly ergodic.*

*Proof.* The proof is inspired by that of [12, Lemma 1.9]. Fix any nonempty word  $u \in \mathcal{A}_p^*$ ; we want to prove that  $N_x(u)$  is uniform Cesàro. Take  $j$  such that  $l = l_j > |u|$ . Further, take arbitrary integers  $1 \leq r < t$  and  $0 \leq s$ ; for abbreviation, write  $N_{st}^{(\cdot)}$  and  $\tau_{st}^{(\cdot)}$  instead of  $N_{x[sl, (s+t)l]}^{(\cdot)}$  and  $\tau_{x[sl, (s+t)l]}^{(\cdot)}$ .

We first prove that

$$(5.4) \quad 0 \leq \tau_{st}^{(1)}(u) - \sum_{v \in \mathcal{A}_p^{rl}} \tau_{st}^{(l)}(v) \cdot \tau_v^{(1)}(u) < \frac{1}{r} + \frac{2r}{t}.$$

To this end, for  $i \in sl + N_{st}^{(1)}(u) \subseteq [sl, (s+t)l]$  put

$$(5.5) \quad \begin{aligned} B_i &= \left\{ h \in \mathbb{N}_0 : [i, i+|u|] \subseteq [hl, (h+r)l] \subseteq [sl, (s+t)l] \right\} \\ &= [s, s+t-r] \cap \left[ \lceil (i+|u|)/l \rceil - r, \lfloor i/l \rfloor \right] \end{aligned}$$

and

$$b_{st} = \sum_{i \in sl + N_{st}^{(1)}(u)} |B_i|.$$

That is,  $b_{st}$  is the number of pairs  $(i, h)$ , where  $i - sl \in N_{st}^{(1)}(u)$  and  $h \in B_i$ . But every such pair  $(i, h)$  corresponds (in a one-to-one way) to a triple  $(v, h', i')$ , where  $v \in \mathcal{A}_p^{rl}$ ,  $h' \in N_{st}^{(l)}(v)$ , and  $i' \in N_v^{(1)}(u)$ ; to see this, put  $v = x[hl, (h+r)l]$ ,  $h' = h$ , and  $i' = i - hl$ . Thus

$$(5.6) \quad b_{st} = \sum_{v \in \mathcal{A}_p^{rl}} |N_{st}^{(l)}(v)| \cdot |N_v^{(1)}(u)|.$$

Further, by (5.5),  $0 \leq |B_i| \leq r$  for every  $i$  and, provided  $(s+r)l \leq i \leq (s+t-r)l$ ,  $|B_i| \geq r-1$ . This gives

$$r \cdot |N_{st}^{(1)}(u)| \geq b_{st} > (r-1) \cdot (|N_{st}^{(1)}(u)| - 2rl)$$

and so

$$(5.7) \quad 0 \leq r \cdot |N_{st}^{(1)}(u)| - b_{st} < |N_{st}^{(1)}(u)| + 2(r-1)rl < (t+2r^2)l.$$

Since  $|N_{st}^{(1)}| = tl\tau_{st}^{(1)}$ ,  $|N_{st}^{(l)}| = t\tau_{st}^{(l)}$ , and  $|N_v^{(1)}| = rl\tau_v^{(1)}$  for  $v \in \mathcal{A}_p^{rl}$ , dividing (5.7) by  $trl$  and using (5.6) gives (5.4).

Now take any  $\varepsilon > 0$ . Let  $j' \geq j$  be such that  $l_{j'}/l > 1/\varepsilon$ ; put  $r = l_{j'}/l$  and  $\varepsilon' = \varepsilon/|\mathcal{A}_p^{r_l}|$ . By the assumption, for every word  $v \in \mathcal{A}_p^{r_l}$  the set  $N_x^{(r_l)}(v)$  is uniform Cesàro; put  $d_v = l \cdot d(N_x^{(r_l)}(v))$ . Thus, by (5.1), we can find  $j'' > j'$  such that  $|\tau_{st}^{(l)}(v) - d_v| < \varepsilon'$  for every  $t \geq l_{j''}/l$  and every  $v \in \mathcal{A}_p^{r_l}$ . We may assume that  $j''$  is so large that  $(2r/t) < \varepsilon$ . Put  $d = \sum_{v \in \mathcal{A}_p^{r_l}} d_v \tau_v^{(1)}(u)$ . Then

$$\left| \sum_{v \in \mathcal{A}_p^{r_l}} \left( \tau_{st}^{(l)}(v) \tau_v^{(1)}(u) - d_v \tau_v^{(1)}(u) \right) \right| < \varepsilon' \sum_{v \in \mathcal{A}_p^{r_l}} \tau_v^{(1)}(u) \leq \varepsilon,$$

and (5.4) gives

$$\left| \tau_{st}^{(1)}(u) - d \right| < 3\varepsilon.$$

This is true for every sufficiently large  $t$  and so, by (5.2), the set  $N_x^{(1)}(u) = \{i : x[i, i + |u|] = u\}$  is uniform Cesàro with density  $d(A) = d$ . Since  $u$  was arbitrary, Lemma 28 yields strict ergodicity of the subshift  $(\text{Orb}_\sigma(x), \sigma)$ .  $\square$

## 6. PROOF OF THEOREM C

In this section we show that Theorem B cannot be generalized to arbitrary dynamical system. We construct a strictly ergodic system for which local correlation entropy of every point is zero, but the topological entropy is positive. The construction is a modification of that from [12, pp. 327–329].

Fix an integer  $p \geq 3$  and take the alphabet  $\mathcal{A} = \mathcal{A}_p = \{0, \dots, p-1\}$ . Recall that  $\mathcal{A}^* = \bigcup_{m \geq 0} \mathcal{A}^m$  denotes the set of all words over  $\mathcal{A}$ . If  $w, v$  are words, their concatenation is denoted by  $wv$ . Further, for a word  $w$  and an integer  $n$ , the concatenation  $ww \dots w$  ( $n$ -times) is denoted by  $w^n$ .

For  $n \geq 1$  denote by  $\mathcal{P}_n$  the set of all permutations  $\pi$  of  $\{1, \dots, n\}$ . Write  $\mathcal{P}_n = \{\pi_1^{(n)}, \dots, \pi_{n!}^{(n)}\}$ , where  $\pi_1^{(n)}$  denotes the identity. For words  $w_1, \dots, w_n \in \mathcal{A}^*$  and  $\pi \in \mathcal{P}_n$  define

$$\pi(w_1, w_2, \dots, w_n) = w_{\pi(1)} w_{\pi(2)} \dots w_{\pi(n)} \in \mathcal{A}^*.$$

Let  $M = \{w_1 < w_2 < \dots < w_n\}$  be an ordered set of words over  $\mathcal{A}$  (the order of  $M$  need not be lexicographical) such that the lengths  $|w_i|$  are the same; denote their common value by  $l(M)$ . For  $r \geq 0$  let  $M^{(r)}$  be the ordered set

$$(6.1) \quad M^{(r)} = \{w_1^{(r)} < \dots < w_n^{(r)}\}, \quad \text{where} \quad w_j^{(r)} = w_1^r \pi_j^n(w_1, w_2, \dots, w_n).$$

Note that the words  $w_j^{(r)}$  are pairwise distinct, the length of every  $w_j^{(r)}$  is  $l(M^{(r)}) = (|M| + r) \cdot l(M)$ , and the cardinality of  $M^{(r)}$  is  $|M^{(r)}| = (|M|)!$ . Further,  $w_1^{(r)}$  starts with  $(r+1)$  copies of  $w_1$ .

Let  $M_1 = \{0 < 1 < \dots < p-1\}$ . Then  $l(M_1) = 1$  and  $|M_1| = p$ . Put  $r_1 = 0$ . If we have defined  $M_j$  and  $r_j$  for  $j \geq 1$ , define  $M_{j+1}$  and  $r_{j+1}$  by

$$(6.2) \quad M_{j+1} = M_j^{(r_j)} \quad \text{and} \quad r_{j+1} = \lceil |M_{j+1}|/l(M_{j+1}) \rceil.$$

For every  $j$  put

$$(6.3) \quad m_j = |M_j|, \quad l_j = l(M_j), \quad \lambda_j = (1/l_j) \log m_j;$$

let  $\bar{w}_j$  denote the first (according to the order of  $M_j$ ) word of  $M_j$ . Note that e.g. for  $j = 1, 2$  we have

$$(6.4) \quad l_1 = 1, \quad m_1 = p, \quad r_1 = 0, \quad l_2 = p, \quad m_2 = p!, \quad r_2 = (p-1)!.$$

Further, for every  $j \geq 1$ ,

$$(6.5) \quad m_{j+1} = m_j! \quad \text{and} \quad l_{j+1} = (m_j + r_j)l_j.$$

Let  $x \in \Sigma_p = \mathcal{A}^{\mathbb{N}_0}$  be the unique sequence such that

$$(6.6) \quad x[0, l_j] = \bar{w}_j;$$

such  $x$  exists since  $\bar{w}_{j+1}$  starts with  $(r_j + 1)$  copies of  $\bar{w}_j$ ;  $x$  is unique since  $l_j = |\bar{w}_j| \nearrow \infty$  by Lemma 30(c) below. Put  $X = \overline{\text{Orb}}_\sigma(x)$ .

The proof of Theorem C goes as follows. First, in Lemmas 31 and 33 we show that the system  $(X, \sigma)$  is strictly ergodic, which will prove (a) of the theorem. The fact that the topological entropy is positive is given in Lemma 35. Finally, local correlation entropies of the system are described in Lemmas 37 and 38. We start by summarizing some of the properties of the constructed sets  $M_j$ .

**Lemma 30.** *The following hold:*

- (a)  $m_j/l_j$  is an even integer provided  $j \geq 2$ , and so  $r_j = m_j/l_j$  (that is, ceiling in (6.2) is unnecessary);
- (b)  $r_j > p$  provided  $j \geq 3$ ;
- (c)  $\lim_j m_j = \lim_j l_j = \lim_j r_j = \infty$ ;
- (d)  $l_{j+1} > pl_j^2$  provided  $j \geq 3$ ;
- (e)  $\sum_{j \geq 4} (1/l_j) < 1/(pl_3^2 - 1)$  and  $l_3 = (p+1)!$ .

*Proof.* (a)–(c) Immediately from the construction we have

$$(6.7) \quad 1 = l_1, \quad p = l_2 < l_3 < \dots, \quad 3 \leq p = m_1 < m_2 < m_3 < \dots$$

and

$$(6.8) \quad l_{j+1} \leq 2m_j l_j \quad \text{for every } j \geq 1$$

(the last inequality follows from  $r_j \leq m_j$ , see (6.2)). Further, we claim that

$$(6.9) \quad l_j \leq m_j - p \quad \text{for every } j \geq 2.$$

Indeed, this is true for  $j = 2$  since  $p \leq p! - p$  (recall that  $p \geq 3$ ). Assume that (6.9) is true for some  $j \geq 2$ . By (6.7),  $2 < l_j \leq m_j - p < m_j - 1$ . Thus  $m_{j+1} = m_j! > m_j(m_j - 1)(m_j - p)2 \geq (m_j - 1) \cdot (2m_j l_j) > (p - 1)l_{j+1}$ . Now  $m_{j+1} - l_{j+1} > (p - 2)l_{j+1} > p$  and (6.9) is true also for  $(j + 1)$ .

By (6.4),  $m_1/l_1 = p$  and  $m_2/l_2 = (p - 1)!$  are integers. Assume now that  $m_j/l_j$  is an integer for some  $j \geq 2$ . Then, by (6.5) and the fact that  $r_j = m_j/l_j$ ,

$$\frac{m_{j+1}}{l_{j+1}} = \frac{m_j!}{(m_j + r_j)l_j} = \frac{(m_j - 1)!}{l_j + 1}.$$

By (6.9), this is an even integer, which is greater than  $(m_j - 2)!$ . Thus (a) is proved and, since  $m_j \geq m_2 = p!$  and  $(m_2 - 2)! \geq m_2 - 2 = p! - 2 \geq 2p - 2 > p$ , also (b) is proved. Further,  $\lim m_j = \lim l_j = \infty$  since these sequences are strictly monotone by (6.7), and  $\lim r_j = \infty$  since, as we have just proved,  $r_{j+1} > (m_j - 2)!$  for  $j \geq 2$ . Thus we have (c).

(d) By (6.5), (a), and (b),  $l_{j+1} = (m_j + r_j)l_j \geq m_j l_j = r_j l_j^2 > pl_j^2$  for  $j \geq 3$ .

(e) A simple induction using (d) gives  $l_{j+3} > p^{2^j-1} l_3^{2^j} \geq (pl_3^2)^j$  for every  $j \geq 1$ . Hence  $\sum_{j \geq 4} (1/l_j) < 1/(pl_3^2 - 1)$ . Since  $l_3 = (p! + (p - 1)!)p = (p + 1)!$  by (6.5) and (6.4), (e) is proved.  $\square$

### 6.1. Strict ergodicity.

**Lemma 31.** *The subshift  $(X, \sigma)$  is minimal.*

*Proof.* Take any word  $u$  which occurs in  $x$ . Then there is  $j$  such that  $u$  occurs in  $\bar{w}_j$ . By the construction,  $u$  occurs in every word from  $M_{j+1}$ . Since  $x$  is a concatenation of words from  $M_{j+1}$ , we have that  $x$  is almost periodic and  $(X, \sigma)$  is minimal.  $\square$

**Lemma 32.** *For every integer  $j \geq 1$  and every word  $v \in M_j$ , the set*

$$N_x^{(l_j)}(v) = \{i \in \mathbb{N}_0 : x[i l_j, (i + 1) l_j] = v\}$$

*is uniform Cesàro with density*

$$d\left(N_x^{(l_j)}(v)\right) = (1/l_{j+1}) \cdot \begin{cases} (r_j + 1) & \text{if } v = \bar{w}_j; \\ 1 & \text{otherwise.} \end{cases}$$

*Proof.* Fix  $j \geq 1$  and  $v \in M_j$ . Take arbitrary  $u \in M_{j+1}$  and write  $u = u_0 u_1 \dots u_{h-1}$ , where  $h = l_{j+1}/l_j$  and  $u_i \in M_j$ . Put  $\bar{\tau}_u(v) = (1/h) \cdot |\{i : u_i = v\}| = \tau_u^{(l_j)}(v)$ . Then, by the construction,

$$\bar{\tau}_u(v) = (1/h) \cdot \begin{cases} (r_j + 1) & \text{if } v = \bar{w}_j; \\ 1 & \text{otherwise} \end{cases}$$



does not depend on  $u$ . This implies that the cardinality of  $N_x^{(l_j)}(v) \cap [il_{j+1}, (i+1)l_{j+1})$  does not depend on  $i$ , and is equal to  $(r_j + 1)$  if  $v = \bar{w}_j$  and to 1 otherwise. By (5.2) and (5.3) the lemma follows.  $\square$

**Lemma 33.** *The subshift  $(X, \sigma)$  is strictly ergodic.*

*Proof.* This immediately follows from Lemma 29. In fact, take any integer  $j \geq 1$  and any  $l_j$ -word  $v$ . If  $v \in M_j$ , then  $N_x^{(l_j)}(v)$  is uniform Cesàro by Lemma 32. Otherwise  $N_x^{(l_j)}(v)$  is empty, so again it is uniform Cesàro. Thus  $(X, \sigma)$  is strictly ergodic by Lemmas 29 and 31.  $\square$

**6.2. Positive topological entropy.** Here we prove that the constructed subshift  $(X, \sigma)$  has positive topological entropy. We start with the crucial fact concerning the sequence  $(\lambda_j)_j$ , the proof of which is a modification of that from [12, Lemma 2.1].

**Lemma 34.** *Let  $p \geq 3$ . Then the sequence  $(\lambda_j)_{j \geq 1}$  is decreasing and the limit of it is positive.*

*Proof.* By (6.5),  $m_{j+1} = m_j! < m_j^{m_j}$ . Hence, using that  $r_j \geq 0$ ,

$$\lambda_{j+1} = \frac{\log m_{j+1}}{l_{j+1}} < \frac{m_j \log m_j}{(m_j + r_j)l_j} \leq \frac{\log m_j}{l_j} = \lambda_j.$$

Thus  $(\lambda_j)_j$  is decreasing. Put  $\lambda = \lim_j \lambda_j$ ; we are going to show that  $\lambda > 0$ . To this end, recall Stirling's formula [21]

$$n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \cdot e^{\delta_n}, \quad \text{where} \quad \frac{1}{12n+1} < \delta_n < \frac{1}{12n} \quad (n \geq 0).$$

For  $n \geq 1$  this yields (using that  $\delta_n > 0$ )

$$(6.10) \quad \log n! > n(\log n - 1).$$

So, by (6.5),  $l_{j+1}\lambda_{j+1} = \log m_{j+1} > m_j(\log m_j - 1)$ . Dividing by  $l_{j+1} = (m_j + r_j)l_j$ , using (6.3) and the facts that  $\log m_j > 0$  (recall that  $\log$  means the natural logarithm) and  $r_j = m_j/l_j > 0$  for  $j \geq 2$ , we have

$$\lambda_{j+1} > \frac{m_j(\log m_j - 1)}{(m_j + r_j)l_j} > \frac{\log m_j}{l_j} - \frac{r_j \log m_j}{m_j l_j} - \frac{1}{l_j}$$

for  $j \geq 2$ . Hence

$$(6.11) \quad \lambda_{j+1} > \lambda_j - \frac{1 + \lambda_j}{l_j} \quad \text{for every } j \geq 1$$

(the inequality is trivial for  $j = 1$  since  $l_1 = 1$  and  $\lambda_2 > 0$ ). By monotonicity of  $(\lambda_j)_j$ , (6.11) implies that, for  $j \geq 1$ ,

$$(6.12) \quad \lambda \geq \lambda_j - (\lambda_j + 1)S_j, \quad \text{where} \quad S_j = \sum_{k \geq j} \frac{1}{l_k}.$$

We want to prove that  $\lambda_j > S_j/(1 - S_j)$  for some  $j$ ; this fact together with (6.12) will imply  $\lambda > 0$ .

Assume first that  $p \geq 4$ . We claim that

$$(6.13) \quad S_2 < \frac{1}{p} + \frac{1}{p(p^2 - 1)} \quad \text{and} \quad \frac{S_2}{1 - S_2} < \frac{1}{p - 2}.$$

Indeed, by Lemma 30(e) and (6.4),  $S_2 < 1/p + 1/l_3 + 1/(pl_3^2 - 1)$  and  $l_3 = (p+1)!$ . Since  $p \geq 4$ ,  $pl_3^2 - 1 > l_3 \geq 2(p-1)p(p+1)$ , we have the first inequality from (6.13). The second one follows from the first and the facts that the map  $x \mapsto x/(1-x)$  is increasing on  $(-\infty, 1)$ , and that  $1/p + 1/[p(p^2 - 1)] < 1$ .

By (6.4) and (6.10),  $\lambda_2 > \log p - 1$ . Since  $\log p > 3/2$  for  $p \geq 5$ ,  $\lambda_2 > 1/2 > S_2/(1 - S_2)$  by (6.13). For  $p = 4$  we have  $\lambda_2 = \log(24)/4$ . Since  $\log(24) > 2$ , we again have  $\lambda_2 > 1/2 > S_2/(1 - S_2)$ . Thus, by (6.12),  $\lambda > 0$  for every  $p \geq 4$ .

It remains to describe the case  $p = 3$ . By (6.4) and (6.5),  $m_3 = 720$ ,  $l_3 = 24$ , and  $\lambda_3 = \log(720)/24$ . Since  $\log(720) > 6$  (use e.g. that  $e < 2.8$ ), we have  $\lambda_3 > (1/4)$ . On the other hand, by Lemma 30(e),  $S_3 < 1/l_3 + 1/(p(l_3^2 - 1)) < 2/l_3 = 1/12$  and  $S_3/(1 - S_3) < 1/11$ . Thus, for  $p = 3$ ,  $\lambda_3 > S_3/(1 - S_3)$  and  $\lambda > 0$  by (6.12).  $\square$

**Lemma 35.** *Let  $p \geq 3$ . Then the topological entropy of the subshift  $(X, \sigma)$  is*

$$h_{\text{top}}(\sigma) = \lim_{j \rightarrow \infty} \lambda_j > 0.$$

*Proof.* Put  $\lambda = \lim_{j \rightarrow \infty} \lambda_j$ ; by Lemma 34 the limit exists and is positive. We prove that  $h_{\text{top}}(\sigma) = \lambda$ ; recall that  $h_{\text{top}}(\sigma) = \lim_n (1/n) \log \theta_n$ , where  $\theta_n$  is the number of  $n$ -words in  $x$  (see e.g. [27, Theorem 7.13]). The inequality  $h_{\text{top}}(\sigma) \geq \lambda$  is trivial, since the number of  $l_j$ -words in  $x$  is larger than or equal to  $m_j$ . To prove the reverse inequality, take any  $l_{j+1}$ -word  $v$  in  $x$ . Since  $l_{j+1} = (m_j + r_j)l_j$ , there are words  $u_1, \dots, u_{m_j+r_j+1}$  from  $M_j$  and an integer  $i \in [0, l_j)$  such that  $v = (u_1 \dots u_{m_j+r_j+1})[i, i + |v|)$ . From this fact it immediately follows that  $\theta_{l_{j+1}} \leq l_j m_j^{m_j+r_j+1}$ . Thus

$$h_{\text{top}}(\sigma) = \lim_{j \rightarrow \infty} \frac{\log \theta_{l_{j+1}}}{l_{j+1}} \leq \lim_{j \rightarrow \infty} \frac{\log l_j + (m_j + r_j + 1) \log m_j}{(m_j + r_j)l_j} = \lambda.$$

□

*Remark 36.* Since the beginning of this section we excluded the case  $p = 2$ . Nevertheless, the construction can be carried over also for such  $p$ . The obtained subshift will be strictly ergodic (by the same reasoning as in Lemmas 31–33). However, the topological entropy will be zero. In fact, by (6.5),

$$l_j = 2 \cdot 3^{j-2}, \quad m_j = 2, \quad r_j = 1 \quad \text{for every } j \geq 2.$$

Hence, for  $j \geq 2$ , the number  $\theta_{l_{j+1}}$  of  $l_{j+1}$  words is less than or equal to  $l_j m_j^{m_j+r_j+1} = 16l_j$  and  $h_{\text{top}}(\sigma) = \lim_j (1/l_{j+1}) \log \theta_{l_{j+1}} = 0$ .

### 6.3. Zero correlation entropy.

**Lemma 37.** *The correlation entropy  $h_{\text{cor}}(\sigma, \mu)$  of the unique invariant measure  $\mu$  of  $(X, \sigma)$  is zero.*

*Proof.* Recall that, by Lemma 28 and the choice of  $x$ ,

$$(6.14) \quad \mu([v]) = \lim_{j \rightarrow \infty} \tau_{\bar{w}_j}(v) \quad \text{for every } v \in \mathcal{A}^*.$$

We start the proof by showing that

$$(6.15) \quad \mu([\bar{w}_j^k]) \geq \frac{r_j - k + 1}{2m_j l_j} \quad \text{for every } j \geq 1 \text{ and } 1 \leq k \leq r_j.$$

To this end, fix any  $j \geq 1$  and  $1 \leq k \leq r_j$ . By the construction, every word  $u$  from  $M_{j+1}$  begins with  $r_j$  copies of  $\bar{w}_j$ . Hence  $u[i l_j, (i+k)l_j) = \bar{w}_j^k$  for every  $0 \leq i \leq (r_j - k)$ . By (6.5) and Lemma 30(a),

$$\tau_u(\bar{w}_j^k) \geq \frac{r_j - k + 1}{l_{j+1}} \geq \frac{r_j - k + 1}{2m_j l_j} \quad \text{for every } u \in M_{j+1}.$$

Now take any  $t > j$ . Since  $\bar{w}_t$  is a concatenation of words from  $M_{j+1}$ , we have  $\tau_{\bar{w}_t}(\bar{w}_j^k) \geq (r_j - k + 1)/(2m_j l_j)$  for every  $u \in M_t$ . By (6.14), this yields (6.15).

Recall the definition (2.4) of  $\tilde{\mu}$ . Take any  $n \in \mathbb{N}$ , put  $w = x[0, n)$ , and find  $j$  such that  $l_j \leq n < l_{j+1}$ ; we may assume that  $j \geq 2$ . Assume first that  $n < (r_j/2)l_j$ ; note that, by Lemma 30(a),  $r_j/2$  is an integer. In this case  $\mu([w]) \geq \mu([\bar{w}_j^{r_j/2}]) \geq (r_j/2)/(2m_j l_j) = 1/(4l_j^2)$  by (6.15) and Lemma 30(a). Thus

$$(6.16) \quad \frac{-\log \tilde{\mu}(n)}{n} \leq \frac{4 \log l_j + 4 \log 2}{l_j}.$$

If  $n \geq (r_j/2)l_j$  then  $\mu([w]) \geq \mu([\bar{w}_{j+1}]) \geq 1/(2l_{j+1}^2) \geq 1/(2^3 m_j^2 l_j^2)$ . Thus

$$(6.17) \quad \frac{-\log \tilde{\mu}(n)}{n} \leq \frac{4 \log m_j + 4 \log l_j + 6 \log 2}{(r_j/2)l_j} = \frac{8\lambda_j}{r_j} + \frac{8 \log l_j + 12 \log 2}{r_j l_j}.$$

Since the right-hand sides of (6.16) and (6.17) converge to zero for  $j \rightarrow \infty$ , we have that  $\lim_n (-1/n) \log \tilde{\mu}(n) = 0$ . So  $h_{\text{cor}}(\sigma, \mu) = 0$  by Lemma 5. □

Since the subshift  $(X, \sigma)$  is strictly ergodic, Proposition 3 immediately implies the following result, which finishes the proof of Theorem C.

**Lemma 38.** *The local correlation entropy  $h_{\text{cor}}(\sigma, y)$  of every  $y \in X$  is zero.*

## ACKNOWLEDGEMENTS

Substantive feedback from Marek Špitalský, Jana Majerová, and Marian Grendár is gratefully acknowledged. This research is an outgrowth of the project “SPAMIA”, MŠ SR-3709/2010-11, supported by the Ministry of Education, Science, Research and Sport of the Slovak Republic, under the heading of the state budget support for research and development. The author also acknowledges support from VEGA 1/0786/15 and APVV-15-0439 grants.

## REFERENCES

- [1] J. Aaronson, R. Burton, H. Dehling, D. Gilat, T. Hill, and B. Weiss. Strong laws for L- and U-statistics. *Trans. Amer. Math. Soc.*, 348(7):2845–2866, 1996.
- [2] H. Broer, F. Takens, and B. Hasselblatt. *Handbook of dynamical systems*, volume 3. Elsevier, 2010.
- [3] P. Collet and J.-P. Eckmann. *Iterated maps on the interval as dynamical systems*. Birkhauser, 1980.
- [4] M. Denker, C. Grillenberger, and K. Sigmund. *Ergodic theory on compact spaces*. Springer, 1976.
- [5] M. Denker and G. Keller. Rigorous statistical procedures for data from dynamical systems. *J. Stat. Phys.*, 44(1-2):67–93, 1986.
- [6] J. P. Eckmann, S. O. Kamphorst, and D. Ruelle. Recurrence plots of dynamical systems. *Europhys. Lett.*, 4(9):973–977, 1987.
- [7] M. Einsiedler and T. Ward. *Ergodic Theory with a view towards Number Theory*. Springer, 2011.
- [8] K. Falconer. *Fractal geometry: mathematical foundations and applications*. John Wiley & Sons, 2004.
- [9] P. Grassberger and I. Procaccia. Characterization of strange attractors. *Phys. Rev. Lett.*, 50(5):346–349, 1983.
- [10] P. Grassberger and I. Procaccia. Measuring the strangeness of strange attractors. *Phys. D*, 9(1-2):189–208, 1983.
- [11] M. Grendár, J. Majerová, and V. Špitalský. Strong laws for recurrence quantification analysis. *Internat. J. Bifur. Chaos*, 23(08):1350147, 2013.
- [12] C. Grillenberger. Constructions of strictly ergodic systems I. Given entropy. *Probab. Theory Related Fields*, 25(4):323–334, 1973.
- [13] F. Hahn and Y. Katznelson. On the entropy of uniquely ergodic transformations. *Trans. Amer. Math. Soc.*, 126(2):335–360, 1967.
- [14] J. Llibre and M. Misiurewicz. Horseshoes, entropy and periods for graph maps. *Topology*, 32(3):649–664, 1993.
- [15] A. Manning and K. Simon. A short existence proof for correlation dimension. *J. Stat. Phys.*, 90(3-4):1047–1049, 1998.
- [16] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths. Recurrence plots for the analysis of complex systems. *Phys. Rep.*, 438(5-6):237–329, 2007.
- [17] J. C. Oxtoby. Ergodic sets. *Bull. Amer. Math. Soc.*, 58(2):116–136, 1952.
- [18] Y. Pesin. Dimension theory in dynamical systems, 1997.
- [19] Y. B. Pesin. On rigorous mathematical definitions of correlation dimension and generalized spectrum for dimensions. *J. Stat. Phys.*, 71(3-4):529–547, 1993.
- [20] Y. B. Pesin and A. Tempelman. Correlation dimension of measures invariant under group action. *Random Comput. Dyn.*, 3(3):137–156, 1995.
- [21] H. Robbins. A remark on Stirling’s formula. *Amer. Math. Monthly*, 62(1):26–29, 1955.
- [22] R. J. Serinko. Ergodic theorems arising in correlation dimension estimation. *J. Stat. Phys.*, 85(1-2):25–40, 1996.
- [23] S. M. Srivastava. *A course on Borel sets*, volume 180. Springer, 2008.
- [24] F. Takens. Invariants related to dimension and entropy. In *Atas do 13 colóquio bras. de mat.*, pages 353–359, 1983.
- [25] F. Takens and E. Verbitskiy. Generalized entropies: Rényi and correlation integral approach. *Nonlinearity*, 11(4):771–782, 1998.
- [26] E. Verbitskiy. *Generalized Entropies in Dynamical Systems, PhD Thesis*. University of Groningen, 2000.
- [27] P. Walters. *An introduction to ergodic theory*, volume 79. Springer, 2000.

- [28] J. P. Zbilut and C. L. Webber. Embeddings and delays as derived from quantification of recurrence plots. *Phys. Lett. A*, 171(3-4):199–203, 1992.

SLOVANET A.S., ZÁHRADNÍČKA 151, BRATISLAVA, SLOVAKIA  
*E-mail address:* `vladimir.spitalsky@slovanet.net`

DEPARTMENT OF MATHEMATICS, FACULTY OF NATURAL SCIENCES, MATEJ BEL UNIVERSITY, TAJOVSKÉHO 40, BANSKÁ  
BYSTRICA, SLOVAKIA  
*E-mail address:* `vladimir.spitalsky@umb.sk`